Neuroeconomics

DECISION MAKING AND THE BRAIN

EDITED BY

PAUL W. GLIMCHER

COLIN F. CAMERER • ERNST FEHR

RUSSELL A. POLDRACK

Academic Press is an imprint of Elsevier
32 Jamestown Road, London NW1 7BY, UK
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA

First edition 2009

Library of Congress Cataloging-in-Publication Data
A catalog record for this book available from the Library of Congress

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-0-12-374176-9

For information on all Academic Press publications
visit our website at www.elsevierdirect.com

Typeset by Charon Tec Ltd., A Macmillan Company. (www.macmillansolutions.com)

Printed and bound in China
08 09 10 11 12   9 8 7 6 5 4 3 2 1

Contributors

Bernard W. Balleine UCLA Department of Psychology, 1285 Franz Hall, Los Angeles, CA 90095-1563, USA

B. Douglas Bernheim Department of Economics, Stanford University, Stanford, CA 94305-6072, USA

Peter Bossaerts Laboratory for Decision Making under Uncertainty, École Polytechnique Fédérale Lausanne (EPFL), Station 5, 1015 Lausanne, Switzerland

Sarah F. Brosnan Department of Psychology, Georgia State University, Atlanta, GA 30302-5010, USA

Julian R. Brown Howard Hughes Medical Institute and Department of Neurobiology, Stanford University School of Medicine, Fairchild Building, Room D200, 299 Campus Drive West, Stanford, CA 94305, USA

Colin F. Camerer Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125, USA

Andrew Caplin Department of Economics, New York University, 269 Mercer Street, New York, NY 10003, USA

M. Keith Chen Yale School of Management, 135 Prospect Street, New Haven, CT 06520-8200, USA

Greg S. Corrado Howard Hughes Medical Institute and Department of Neurobiology, Stanford University School of Medicine, Fairchild Building, Room D200, 229 Campus Drive West, Stanford, CA 94305, USA

Antonio Damasio Brain and Creativity Institute, University of Southern California, Los Angeles, CA 90089-2520, USA

Nathaniel D. Daw Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, USA

Peter Dayan Gatsby Computational Neuroscience Unit, Alexandra House, 17 Queen Square, London WC1N 3AR, UK

Mark Dean Department of Economics, New York University, 19 West 4th Street, New York, NY 1003, USA

Mauricio R. Delgado Department of Psychology, Rutgers University, 101 Warren Street, Smith Hall, Newyork, NJ 07102, USA

Michael Dorris Department of Physiology, Queen's University, 18 Stuart Street, Botterell Hall, Kingston, ON K7L3N6, Canada

Kenji Doya Neural Computation Unit, Okinawa Institute of Science and Technology, 12–22 Suzaki, Uruma, Okinawa 904-2234, Japan

Ernst Fehr Institute for Empirical Research in Economics, University of Zürich, Blümlisalpstrasse 10, CH-8006 Zürich, Switzerland

Craig R. Fox UCLA Anderson School and Department of Psychology, 110 Westwood Plaza #D511, Los Angeles, CA 90095-1481, USA

Charles R. Gallistel Rutgers Center for Cognitive Science, Rutgers University, Psychology Building Addition, Busch Campus, 152 Frelinghuysen Road, Piscataway, NJ 08854-8020, USA

Paul W. Glimcher Center for Neuroeconomics, New York University, 6 Washington Place, New York, NY 10013, USA

William T. Harbaugh Department of Economics, University of Oregon, Eugene, Oregon 97403-1285, USA

Daniel Houser Interdisciplinary Center for Economic Science (ICES), 4400 University Drive, Fairfax, VA 22030, USA

Ming Hsu Department of Economics, University of Illinois at Urbana-Champaign, 405 North Mathews Avenue, Urbana, IL 61801, USA

Eric J. Johnson Columbia Business School, Columbia University, Uris Hall, 3022 Broadway, New York, NY 10027, USA

Daniel Kahneman Center for Health and Well-Being, and Woodrow Wilson School of Public and International Affairs, 322 Wallace Hall, Princeton University, Princeton, NJ 08544, USA

Minoru Kimura Division of Neurophysiology, Kyoto Prefectural University of Medicine, Kawaramachi-Hirokoji, Makigyo-ku, Kyoto 602-8566, Japan

Brian Knutson Department of Psychology and Neuroscience, Stanford University, 470 Jordan Hall, Stanford, CA 94305-2130, USA

Michael S. Landy Department of Psychology and Center for Neural Science, New York University, 6 Washington Place, New York, NY 10003, USA

Daeyeol Lee Department of Neurobiology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510, USA

Laurence T. Maloney Department of Psychology and Center for Neural Science, New York University, 6 Washington Place, New York, NY 10003, USA

Ulrich Mayr Department of Psychology, University of Oregon, Eugene, Oregon 97403-1227, USA

Kevin McCabe Department of Economics, George Mason University, Fairfax, VA 22030-4444, USA

P. Read Montague Department of Neuroscience, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX 7703, USA

William T. Newsome Howard Hughes Medical Institute and Department of Neurobiology, Stanford University School of Medicine, Fairchild Building, Room D209, 299 Campus Drive West, Stanford, CA 94305, USA

Yael Niv Center for the Study of Brain, Mind and Behaviour, Department of Psychology, and Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08544, USA

John P. O'Doherty Division of Humanities and Social Sciences, California Institute of Technology, 1200 E California Boulevard, Pasadena, CA 91125, USA

Camillo Padoa-Schioppa Department of Neurobiology, Harvard Medical School, 220 Longwood Avenue, Boston, MA 02115, USA

Elizabeth A. Phelps Center for Neuroeconomics, New York University, 6 Washington Place, New York, NY 10003, USA

Paul E. M. Phillips Department of Psychiatry & Behavioral Sciences and Department of Pharmacology, University of Washington, Seattle, WA 98195-6560, USA

Michael Platt Department of Neurobiology, Duke University Medical Center, 427E Bryan Research Building, Durham, NC 27710, USA

Russell A. Poldrack UCLA Department of Psychology, 1285 Franz Hall, Los Angeles, CA 90095-1563, USA

Kerstin Preuschoff University of Zürich, Blümlisalpstrasse 10, CH-8006 Zürich, Switzerland

Antonio Rangel Division of Humanities and Social Sciences, California Institute of Technology (Caltech), HSS 228-77, Pasadena, CA 91125-7700, USA

Aldo Rustichini Department of Economics, University of Minnesota, 1035 Heller Hall, 271 19th Avenue South, Minneapolis, MN 55455, USA

Alan Sanfey Department of Psychology, University of Arizona, Tucson, AZ 85719, USA

Laurie R. Santos Department of Psychology, Yale University, New Haven, CT 06510, USA

Wolfram Schultz Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK

Ben Seymour Wellcome Department of Imaging Neuroscience, University College London, London WC1N 3BG, UK

Joan B. Silk UCLA Department of Anthropology, 341 Haines Hall, Los Angeles, CA 90095-1553, USA

Tania Singer Center for Social Neuroscience and Neuroeconomics, University of Zürich, Blümlisalpstrasse 10, CH-8006 Zürich, Switzerland

Vernon L. Smith Department of Economics, MSN 3G4, George Mason University, Fairfax, VA 22030-4444, USA

Leo P. Sugrue Howard Hughes Medical Institute and Department of Neurobiology, Stanford University School of Medicine, Fairchild Building, Room D200, Stanford, CA 94305, USA

Dharol Tankersley Brain Imaging and Analysis Center, Duke University Medical Center, Durham, NC 27710, USA

Julia Trommershäuser Department of Psychology, Giessen University, Otto-Behaghel-Str. 10F, 35394 Giessen, Germany

Xiao-Jing Wang Department of Neurobiology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510, USA

Elke U. Weber Department of Psychology, Columbia University, 1190 Amsterdam Avenue, New York, NY 10027, USA

Preface

Over the past decade there has been a tremendous growth in both scholarly and popular interest at the intersection of neuroscience, economics, and psychology. Fifteen years ago less than four academic papers were published a year that were tagged with both "brain" and "decision making" as keywords. Today, almost 200 are published each year and that number doubles approximately every three years. What the field has lacked until now, however, is a comprehensive source for academic scholars that provides a complete survey of the field at a technical level. It is our hope that this volume will fill that gap.

USING THE BOOK AS A HANDBOOK

As editors, we see this book as filling three specific niches. First, we see the book as a "Handbook of Neuroeconomics". A volume that can be picked up by a practicing economist, psychologist, or neuroscientist from which he or she can gain a fairly intimate understanding of the accomplishments and challenges in Neuroeconomics today. For this reason, each chapter has been written to stand alone as an independent contribution. For a reader looking to gain a deeper understanding of one or more of the subareas of this field, the chapters can be read in any order.

USING THE BOOK AS A TEXTBOOK

Second, we see the book as a graduate (or advanced undergraduate) textbook appropriate for use in a seminar course on Neuroeconomics. Our goal in designing the book and editing the chapters was to create a text that beginning graduate students in any department would find both readable and informative. Our goal was for each chapter to both provide necessary background information for interdisciplinary students and offer sufficient depth for experts. To achieve that end, we have worked with the authors to minimize the use of technical vocabulary (wherever possible) and have structured the book into five sections. Looking through these sections the reader will note, as an added feature, that the first 1–3 chapters of each section fill two roles. They both provide critical pedagogical background material for interdisciplinary study and survey an important advance within Neuroeconomics. Teachers using the book as a text are urged to consider this feature when making assignments. Our hope is that reading these introductory chapters will provide students (and faculty) from any discipline with enough background material to understand the critical issues in the field and the more technical chapters that follow.

For this reason, we suggest that the order of section presentation in a classroom be customized according to the field in which the students are experts. For psychologists of Judgment and Decision Making, section two will contain the most familiar material and this might serve as an excellent starting point. For social psychologists, section three might be an appropriate starting point. Economists will find section one a comfortable place to start just as neurobiologists will find sections four and five particularly familiar. Alternatively, of course, the book can be read from cover-to-cover and the result of that approach should be a solid starting point for future work in all of the parent disciplines from which Neuroeconomics is drawn.

We do recognize, however, that many students will find much in this volume that is very new to them. For those students, we specifically recommend a number of companion texts that we have used with our own students. For economists with no starting knowledge of the brain whatsoever it may be helpful to read Rosenzweig, Breedlove, and Watson's *Biological Psychology*. For neuroscientists and psychologists new to the study of decision-making we suggest as a companion Scott Plous' award winning book: *The Psychology of Judgment and Decision Making*. For neuroscientists and psychologists with strong mathematical backgrounds (and those particularly interested in the neoclassical tradition as it applies to microeconomics) we suggest either David Krep's *A Course in Microeconomic Theory* or Mas-Colell, Whinston, and Green's *Microeconomic Theory*. For those same readers interested in game theory we suggest either Fudenberg and Tirole's *Game Theory* or Osborne and Rubenstein's *A Course in Game Theory*.

THE BOOK AS A TIME CAPSULE

Finally, we see the book as a kind of time capsule that documents the field of Neuroeconomics just as it is beginning. Inside the covers of this book are most of the important trends we can identify today. It is with excitement that all four of us look forward to leafing through the book in a decade to two, when the dramatic insights of this early period can be seen through a longer lens.

Acknowledgements

In closing, we absolutely must thank the many people whose hard work made this volume possible. First and foremost among them are Samanta Shaw and Maggie Grantner. Samanta Shaw served as the book's good mother, she labored tirelessly, shepherding each chapter (and each author) through submission and revision. It is she more than anyone else who made the book a reality and she has our undying thanks. Maggie Grantner served as the book's godmother. Founding administrative director of the Society for Neuroeconomics, she contributed to the volume not just as an administrator but as a thoughtful and scholarly critic who oversaw every stage of production. We would also like to express our gratitude to Johannes Menzel, our editor at Academic Press.

Paul W. Glimcher
Colin F. Camerer
Ernst Fehr
Russell A. Poldrack

CHAPTER

1

Introduction: A Brief History of Neuroeconomics

Paul W. Glimcher, Colin F. Camerer, Ernst Fehr, and Russell A. Poldrack

OUTLINE

Over the first decade of its existence, neuroeconomics has engendered raucous debates of two kinds. First, scholars within each of its parent disciplines have argued over whether this synthetic field offers benefits to their particular parent discipline. Second, scholars within the emerging field itself have argued over what form neuroeconomics should take. To understand these debates, however, a reader must understand both the intellectual sources of neuroeconomics and the backgrounds and methods of practicing neuroeconomists.

Neuroeconomics has its origins in two places; in events following the neoclassical economic revolution of the 1930s, and in the birth of cognitive neuroscience during the 1990s. We therefore begin this brief history with a review of the neoclassical revolution and the birth of cognitive neuroscience.

NEOCLASSICAL ECONOMICS

The birth of economics is often traced to Adam Smith's publication of *The Wealth of Nations* in 1776. With this publication began the classical period of economic theory. Smith described a number of phenomena critical for understanding choice behavior and the aggregation of choices into market activity. These were, in essence, psychological insights. They were relatively *ad hoc* rules that explained how features of the environment influenced the behavior of a nation of consumers and producers.

What followed the classical period was an interval during which economic theory became very heterogenous. A number of competing schools with different approaches developed. Many economists of the time

(Edgeworth, Ramsey, Fisher) dreamed about tools to infer value from physical signals, through a "hedonimeter" for example, but these early neuroeconomists did not have such tools (Colander, 2008).

One school of thought, due to John Maynard Keynes, was that regularities in consumer behavior could (among other things) provide a basis for fiscal policy to manage economic fluctuations. Many elements in Keynes' theory, such as the "propensity to consume" or entrepreneurs' "animal spirits" that influence their investment decisions, were based on psychological concepts. This framework dominated United States' fiscal policy until the 1960s.

Beginning in the 1930s, a group of economists – most famously, Samuelson, Arrow, and Debreu – began to investigate the mathematical structure of consumer choice and behavior in markets (see, for example, Samuelson, 1938). Rather than simply building models that incorporated a set of parameters that might, on *a priori* psychological grounds, be predictive of choice behavior, this group of theorists began to investigate what mathematical structure of choices might result from simple, more "primitive," assumptions on preferences. Many of these models (and the style of modeling that followed) had a strong normative flavor, in the sense that attention was most immediately focused on idealized choices and efficient allocation of resources; as opposed to necessarily seeking to describe how people choose (as psychologists do) and how markets work.

To better understand this approach, consider what is probably the first and most important of these simple models: the *Weak Axiom of Revealed Preference* (WARP). WARP was developed in the 1930s by Paul Samuelson, who founded the revealed preference approach that was the heart of the neoclassical revolution. Samuelson proposed that if a consumer making a choice between an apple and an orange selects an apple, he *reveals a preference* for apples. If we assume only that this means he *prefers* (preference is here a stable internal property that economists did not hope to measure directly) apples to oranges, what can we say about his future behavior? Can we say anything at all?

What Samuelson and later authors showed mathematically was that even simple assumptions about binary choices, revealing stable (weak) preferences, could have powerful implications. An extension of the WARP axiom called GARP (the "generalized" axiom of revealed preference, Houthakker, 1950) posits that if apples are revealed preferred to oranges, and oranges are revealed preferred to peaches, then apples are "indirectly" revealed preferred to peaches (and similarly for longer chains of indirect revelation). If GARP holds for binary choices among pairs of objects, then some choices can be used to make predictions about the relative desirability of pairs of objects that have never been directly compared by the consumer. Consider a situation in which a consumer chooses an *apple* over an *orange* and then an *orange* over a *peach*. If the assumption of GARP is correct, then this consumer must not choose a *peach* over an *apple* even if this is a behavior we have never observed before.

The revealed preference approach thus starts from a set of assumptions called axioms which encapsulate a theory of some kind (often a very limited one) in formal language. The theory tells us what a series of observed choices implies about intermediate variables such as utilities (and, in more developed versions of the theory, subjective beliefs about random events). The poetry in the approach (what distinguishes a beautiful theory from an ugly one) is embodied in the simplicity of the axioms, and the degree to which surprisingly simple axioms make sharp predictions about what kind of choice patterns should and should not be observed. Finally, it is critical to note that what the theory predicts is which new choices could possibly follow from an observed set of previous choices (including choices that respond to policy and other changes in the environment, such as responses to changes in prices, taxes, or incomes). The theories do not predict intermediate variables; they use them as tools. What revealed preference theories predict is choice. It is the only goal, the only reason for being, for these theories.

What followed the development of WARP were a series of additional theorems of this type which extended the scope of revealed-preference theory to choices with uncertain outcomes whose likelihoods are known (von Neumann and Morgenstern's expected utility theory, EU) or subjective (or "personal," in Savage's subjective EU theory), and in which outcomes may be spread over time (discounted utility theory) (see Chapter 3 for more details). What is most interesting about these theories is that they demonstrate, amongst other things, that a chooser who obeys these axioms must behave both "as if" he has a continuous utility function that relates the subjective value of any gain to its objective value and "as if" his actions were aimed at maximizing total obtained utility. In their seminal book von Neumann and Morgenstern also laid the foundations for much of *game theory*, which they saw as a special problem in utility theory, in which outcomes are generated by the choices of many players (von Neumann and Morgenstern, 1944).

At the end of this period, neoclassical economics seemed incredibly powerful. Starting with as few as one and as many as four simple assumptions which fully described a new theory the neoclassicists developed a framework for thinking about and predicting choice. These theories of consumer choice would later form the basis for the demand part of the Arrow-Debreu theory of competitive "general" equilibrium, a system in which prices and quantities of all goods were determined simultaneously by matching supply and demand. This is an important tool because it enables the modeler to anticipate *all* consequences of a policy change – for example, imposing a luxury tax on yachts might increase crime in a shipbuilding town because of a rise in unemployment there. This sort of analysis is unique to economics, and partly explains the broad influence of economics in regulation and policy-making.

It cannot be emphasized enough how much the revealed-preference view suppressed interest in the psychological nature of preference, because clever axiomatic systems could be used to infer properties of unobservable preference from observable choice (Bruni and Sugden, 2007). Before the neoclassical revolution, Pareto noted in 1897 that

> It is an empirical fact that the natural sciences have progressed only when they have taken secondary principles as their point of departure, instead of trying to discover the essence of things. … Pure political economy has therefore a great interest in relying as little as possible on the domain of psychology.
>
> (Quoted in Busino, 1964: xxiv)

Later, in the 1950s, Milton Friedman wrote an influential book, *The Methodology of Positive Economics*. Friedman argued that assumptions underlying a prediction about market behavior could be wrong, but the prediction could be approximately true. For example, even if a monopolist seller does not sit down with a piece of paper and figure out what price maximizes total profit, monopoly prices might evolve "as if" such a calculation has been made (perhaps due to selection pressures within or between firms). Friedman's argument licensed economists to ignore evidence of when economic agents violate rational-choice principles (evidence that typically comes from experiments that test the individual choice principles most clearly), a prejudice that is still widespread in economics.

What happened next is critical for understanding where neuroeconomics arose. In 1953, the French economist Maurice Allais designed a series of pairwise choices which led to reliable patterns of revealed preference that violated the central "independence" axiom of expected utility theory. Allais unveiled his pattern, later called the "Allais paradox," at a conference in France at which many participants, including Savage, made choices which violated their own theories during an informal lunch. (Savage allegedly blamed the lunchtime wine.)

A few years after Allais' example, Daniel Ellsberg (1961) presented a famous paradox suggesting that the "ambiguity" (Ellsberg's term) or "weight of evidence" (Keynes' term) supporting a judgment of event likelihood could influence choices, violating one of Savage's key axioms. The Allais and Ellsberg paradoxes raised the possibility that the specific functional forms of EU and subjective EU implied by simple axioms of preference were generally wrong. More importantly, the paradoxes invited mathematical exploration (which only came to fruition in the 1980s) about how weaker systems of axioms might generalize EU and SEU. The goal of these new theories was to accommodate the paradoxical behavior in a way that is both psychologically plausible and formally sharp (i.e., which does not predict that any pattern of choices is possible, and could therefore conceivably be falsified by new paradoxes).

One immediate response to this set of observations was to argue that the neoclassical models worked, but only under some limited circumstances – a fact which many of the neoclassicists were happy to concede (for example, Morgenstern said "the probabilities used must be within certain plausible ranges and not go to .01 or even less to .001"). Surely axioms might also be violated if the details of the options being analyzed were too complicated for the chooser to understand, or if the chooser was overwhelmed with too many choices. Observed violations could then be seen as a way to map out boundary conditions – a specification of the kinds of problems that lay outside the limits of the neoclassical framework's range of applicability.

Another approach was Herbert Simon's suggestion that rationality is computationally bounded, and that much could be learned by understanding "procedural rationality." As a major contributor to cognitive science, Simon clearly had in mind theories of choice which posited particular procedures, and suggested that the way forward was to understand choice procedures empirically, perhaps in the form of algorithms (of which "always choose the object with the highest utility" is one extreme and computationally demanding procedure).

A sweeping and constructive view emerged from the work of Daniel Kahneman and Amos Tversky (1979) in the late 1970s and 1980s, and other psychologists interested in judgment and decision making whose interests intersected with choice theory. What Kahneman, Tversky, and others showed in a series of remarkable experimental examples was that the range of phenomena that fell outside classical expected utility theory was even broader than Allais' and Ellsberg's examples had suggested.

These psychologists studying the foundations of economic choice found many common choice behaviors – typically easily replicated in experiments – that falsified one or more of the axioms of expected utility theory and which seemed to conflict with fundamental axioms of choice. For example, some of their experimental demonstrations showed effects of "framing," attacking the implicit axiom of "description invariance" – the idea that choices among objects should not depend on how they are described.

These experiments thus led many scholars, particularly psychologists and economists who had become interested in decision making through the work of Kahneman and Tversky, to conclude that empirical critiques of the simple axiomatic approaches, in the form of counterexamples, could lead to more general axiomatic systems that were more sensibly rooted in principles of psychology.

This group of psychologists and economists, who began to call themselves *behavioral economists*, argued that evidence and ideas from psychology could improve the model of human behavior inherited from neoclassical economics. In one useful definition, behavioral economics proposes models of limits on rational calculation, willpower, and self-interest, and seeks to codify those limits formally and explore their empirical implications using mathematical theory, experimental data, and analysis of field data.

In the realm of risky choice, Kahneman and Tversky modified expected utility to incorporate a psychophysical idea of reference-dependence – valuation of outcomes depends on a point of reference, just as sensations of heat depend on previous temperature – along with a regressive non-linear transformation of objective probability. (Details of prospect theory are reviewed in Chapter 11.) Another component of the behavioral program was the idea that statistical intuitions might be guided by *heuristics*, which could be inferred empirically by observing choice under a broad range of circumstances. Heuristics were believed to provide a potential basis for a future theory of choice (Gilovich *et al.*, 2002). A third direction is theories of social preference – how people value choices when those choices impact the values of other people (see Chapter 15). The goal is eventually to have mathematical systems that embody choice heuristics and specific types of social preference which explain empirical facts but also make sharp predictions. Development of these theories, and tests with both experimental and field data, are now the frontiers of modern behavioral economics.

An obvious conflict developed (and continues to cause healthy debate) between the behavioral economists, who were attempting to piece together empirically disciplined theories, and the neoclassicists, who were arguing for a simpler global theory, typically guided by the idea that normative theory is a privileged starting point. The difference in approaches spilled across methodological boundaries too. The influence of ideas from behavioral economics roughly coincided with a rise in interest among economists such as Charles Plott, Vernon Smith and colleagues in conducting carefully controlled experiments on economics systems (see, for example, Smith, 1976). The *experimental economists* began with the viewpoint that economic principles should apply everywhere (as principles in natural and physical sciences are presumed to); their view was that when theories fail in simple environments, those failures raise doubt about whether they are likely to work in more complex environments. However, the overlap between behavioral economics and experimental economics is far from complete. Behavioral economics is based on the presumption that incorporating psychological *principles* will improve economic analysis, while experimental economics presumes that incorporating psychological *methods* (highly controlled experiments) will improve the testing of economic theory.

In any case, the neoclassical school had a clear theory and sharp predictions, but the behavioral economists continued to falsify elements of that theory with compelling empirical examples. Neuroeconomics emerged from within behavioral and experimental economics because behavioral economists often proposed theories that could be thought of as algorithms regarding how information was processed, and the choices that resulted from that information-processing. A natural step in testing these theories was simultaneously to gather information on the details of both information processing and associated choices. If information processing could be hypothesized in terms of neural activity, then neural measures could be used (along with coarser measures like eyetracking of information that choosers attend to) to test theories as simultaneous restrictions on what information is processed, how that processing works in the brain, and the choices that result. Neuroscientific tools provide further predictions in tests with lesion-patient behavior, and transcranial magnetic stimulation (TMS) which should (in theory) change choices if TMS disrupts an area that is necessary to producing certain kinds of choices. An important backdrop to this development is that economic theorists are extremely clever at inventing multiple systems of axioms which can explain the same patterns of choices. By definition, choices alone provide a limited way to distinguish theories in the face of rapid production of alternative theories. Forcing theories to commit to predictions about underlying neural activity therefore provides a powerful way to adjudicate among theories.

COGNITIVE NEUROSCIENCE

Like economics, the history of the neuroscientific study of behavior also reflects an interaction between two approaches – in this case, a neurological approach and a physiological approach. In the standard neurological approach of the last century, human patients or experimental animals with brain lesions were studied in a range of behavioral tasks. The behavioral deficits of the subjects were then correlated with their neurological injuries and the correlation used to infer function. The classic example of this is probably the work of the British neurologist David Ferrier (1878), who demonstrated that destruction of the precentral gyrus of the cortex led to quite precise deficits in movement generation. What marks many of these studies during the classical period in neurology is that they often focused on damage to either sensory systems or movement control systems. The reason for this should be obvious; the sensory stimuli presented to a subject are easy to control and quantify – they are *observables* in the economic sense of the word. The same is true for movements that we instruct a subject to produce. Movements are directly observable and easily quantified. In contrast, mental state is much more elusive. Although there has for centuries been clear evidence that neurological damage influences mental state, relating damage to mental state is difficult specifically because mental state is not directly observable. Indeed, relating mental state to neurological damage requires some kind of theory (often a global one), and it was this theory that was largely absent during the classical period in neurology.

In contrast to the neurological approach, the physiological approach to the study of the brain involves correlating direct measurements of biological state, such as the firing of action potentials in neurons, changes in blood flow, and changes in neurotransmitters, with events in the outside world. During the classical period this more precise set of methodological tools was extremely powerful for elucidating basic features of nervous function, but was extremely limited in its applicability to complex mental states. Initially this limitation arose from a methodological constraint. Physiological measurements are invasive and often destructive. This limits their use in animals and, in the classical period, in anesthetized animals. The result was an almost complete restriction of physiological approaches during the classical period to the study of sensory encoding in the nervous system.

A number of critical advances during the period from the 1960s to the 1980s, however, led to both a broadening of these approaches and, later, a fusion of these two approaches. Within the domain of neurology, models from psychology began to be used to understand the relationship between brain and behavior. Although the classes of models that were explored were highly heterogeneous and often not very quantitative, these early steps made it possible to study mental state, at least in a limited way. Within the physiological tradition, technical advances that led to the development of humane methods made it possible to make measurements in awake, behaving animals, also opening the way to the study of mental state, this time in animals.

What followed was a period in which a heterogeneous group of scholars began to develop models of mental processes and then correlate intermediate variables in these models with either physiological measurements or lesion-induced deficits. However, these scholars faced two very significant problems. First, there was a surplus of models. Dozens of related models could often account for the same phenomena, and it was hard to discriminate between these models. Second, there was a paucity of data. Physiological experiments are notoriously difficult and slow, and although they yield precise data they do so at an agonizingly slow rate. Neurological experiments (at least in humans) move more quickly but are less precise, because the researcher does not have control over the placement of lesions.

It was the resolution of these two problems, or attempts to resolve them, that was at the heart of the cognitive neuroscientific revolution. In describing that revolution, we focus on the study of decision making. This was by no means a central element in the cognitive neuroscientific revolution, but it forms the central piece for understanding the source of neuroeconomics in the neuroscientific community.

The lack of a clear global theory was first engaged seriously by the importation of signal detection theory into the physiological tradition. Signal detection theory (Green and Swets, 1966) is a normative theory of signal categorization broadly used in the study of human perception. The critical innovation that revolutionized the physiological study of cognitive phenomena was the use of this normative theory to relate neuronal activity directly to behavior.

In the late 1980s, William Newsome and J. Anthony Movshon (see, for example, Newsome *et al.*, 1989) began work on an effort to relate the activity of neurons in the middle temporal area of visual cortex (Area MT) to decisions made by monkeys in the domain of perceptual categorization. In those experiments, thirsty monkeys had to evaluate an ambiguous visual signal which indicated which of two actions would yield a fluid reward. What the experiments demonstrated was that the firing rates of single neurons in this area, which were hypothesized to encode the perceptual signal being directly evaluated by the monkeys in their decision making, could be used to predict the patterns of stochastic choice produced by the animals in response to the noisy sensory signals. This was a landmark event in neuroscience, because it provided the first really clear demonstration of a correlation between neuronal activity and stochastic choice. Following Newsome's suggestion, this class of correlation came to be known as a *psychometric–neurometric match* – the behavioral measurement being referred to as psychometric and the matching neuronal measurement as neurometric.

This was also a landmark event in the neural study of decision making, because it was the first successful attempt to predict decisions from single neuron activity. However, it was also controversial. Parallel studies in areas believed to control movement generation (Glimcher and Sparks, 1992) seemed not to be as easily amenable to a signal-detection based analysis (Sparks, 1999; Glimcher, 2003). This led to a long-lasting debate in the early and mid-1990s regarding whether theories such as signal detection would prove adequate for the wholesale study of decision making.

The neurological tradition had gained its first glimpses into the effects of brain damage on decision making in 1848, in the case of Phineas Gage (Macmillan, 2002). After his brain was penetrated by a steel rod, Gage exhibited a drastic change in personality and decision-making ability. The systematic study of decision-making deficits following brain damage was initially undertaken, in the 1990s, by Antonio Damasio, Antoine Bechara, and their colleagues (see, for example, Bechara *et al.*, 1994), who began examining decision making under risk in a card-sorting experiment. Their work related damage to frontal cortical areas with specific elements of an emotion-based theory of decision making which, though not normative like signal detection theory, was widely influential. The interest in decision making that this work sparked in the neurological community was particularly opportune, because at this time the stage was being set for combining a new kind of physiological measurement with behavioral studies in humans.

A better understanding of the relation between mental and neural function in humans awaited the development of methods to image human brain activity non-invasively. Early work by Roland, Raichle, and others had used positron emission tomography (PET) to image the neural correlates to mental function, but this method was limited in its application owing to the need for radioactive tracers. In 1992, three groups (Bandettini *et al.*, 1992; Kwong *et al.*, 1992; Ogawa *et al.*, 1992) simultaneously published the first results using functional magnetic resonance imaging (fMRI) to image brain activity non-invasively – a development that opened the door for direct imaging of brain activity while humans engaged in cognitive tasks. This was a critical event, because it meant that a technique was available for the rapid (if crude) direct measurement of neural state in humans. Owing to the wide availability of MRI and the safety of the method, the use of fMRI for functional imaging of human cognitive processes has grown exponentially. Perhaps because of the visually compelling nature of the results, showing brain areas "lighting up," this work became highly influential not just in the neuroscientific and psychological communities but also beyond. The result was that scholars in many disciplines began to consider the possibilities of measuring the brain activity of humans during decision making. The challenge was that there was no clear theoretical tool for organizing this huge amount of information.

SETTING THE STAGE FOR NEUROECONOMICS

By the late 1990s, several converging trends had set the stage for the birth of neuroeconomics. Within economics and the psychology of judgment and decision making, a critical tension had emerged between the neoclassical/revealed preference school and the behavioral school. The revealed-preference theorists had an elegant axiomatic model of human choice which had been revealed to be only crudely predictive of human behavior, and for which it was easy to produce counterexamples. Revealed-preference theorists responded to this challenge by both tinkering with the model to improve it and challenging the significance of many of the existing behavioral economic experiments (relying on the Friedman "F-twist" – that predictions based on axioms might be approximately true even if the axioms are wrong).

The behavioral economists, in contrast, responded to this challenge by looking for alternative mathematical theories and different types of data to test those theories – theories which they saw as being claims about both computational processes and choices. Their goal was to provide an alternative theoretical approach for predicting behavior and a methodology for testing those theories. This is an approach that requires good theories that predict both choices and "non-choice" data. The appropriate form for such an alternative theory has, however, been hotly debated. One approach to developing such a theory derives from the great progress economics has made towards understanding the interaction of two agent systems in the external world – for example, understanding the interactions of firms and the workers they hire. This pre-existing mathematical facility with two-agent models aligned naturally with an interest among psychologists in what are known as “dual-process” models. If, as some behavioral economists have argued, the goal is to minimally complicate the standard models from economics, then going from a single agent maximizing a unifying “utility” to two independent agents (or processes) interacting might be a useful strategy. This strategy forms one of the principle alternative theoretical approaches that gave birth to neuroeconomics. The appeal of the dual-process model for economists is that when inefficient choice behaviors are observed in humans, these can be viewed as the result of the two (or more) independent agents being locked in a bad equilibrium by their own self-interests. Of course, other scholars within behavioral economics have suggested other approaches that also have neuroeconomic implications. A view from evolutionary psychology that may serve as another example is that encapsulated models execute heuristics that are specially adapted to evolutionarily selected tasks (see, for example, Gigerenzer *et al.*, 2000). These models have something to say about the tradeoff between efficient choice and computational complexity, which might be used to generate hypotheses about brain processes (and cross-species comparisons).

Within much of neuroscience, and that fraction of cognitive psychology closely allied with animal studies of choice, a different tension was simultaneously being felt as these multiple agent and heuristic models were evolving in behavioral economics. It was clear that both those physiologists interested in single neuron studies of decision making and those cognitive neuroscientists closely allied to them were interested in describing the algorithmic mechanisms of choice. Their goal was to describe the neurobiological hardware that supported choice behavior in situations ranging from perceptual decision making to the expression of more complicated preferences. What they lacked was an overarching theoretical framework for placing their neural measurements into context. Newsome and his colleagues had argued that the standard mathematical tool for understanding sensory categorization – signal detection theory – could serve that role, but many remained skeptical that this approach could be sufficiently generalized. What that naturally led to was the suggestion, by Glimcher and his colleagues, that the neoclassical/revealed preference framework might prove a useful theoretical tool for neuroscience. What followed was the rapid introduction to the neuroscientific literature of such concepts as expected value and expected utility.

TWO TRENDS, ONE GOAL

The birth of neuroeconomics, then, grew from a number of related factors that simultaneously influenced what were basically two separate communities, albeit with a significant overlap. A group of behavioral economists and cognitive psychologists looked towards functional brain-imaging as a tool to both test and develop alternatives to neoclassical/revealed preference theories (especially when too many theories chased too few data using choices as the only class of data). A group of physiologists and cognitive neuroscientists looked towards economic theory as a tool to test and develop algorithmic models of the neural hardware for choice. The result was an interesting split that persists in neuroeconomics today – and of which there is evidence in this volume.

The result is that the two communities, one predominantly (although not exclusively) neuroscientific and the other predominantly (although not exclusively) behavioral economic, thus approached a union from two very different directions. Both, however, promoted an approach that was controversial within their parent disciplines. Many neurobiologists outside the emerging neuroeconomic community argued that the complex normative models of economics would be of little value for understanding the behavior of real humans and animals. Many economists, particularly hardcore neoclassicists, argued that algorithmic-level studies of decision making were unlikely to improve the predictive power of the revealed-preference approach.

Despite these challenges, the actual growth of neuroeconomics during the late 1990s and early 2000s was explosive. The converging group of like-minded economists, neuroscientists, and cognitive psychologists quickly generated a set of meetings and conferences that fostered a growing sense of interdisciplinary collaboration. Probably the first of these interdisciplinary interactions was held in 1997 at Carnegie-Mellon University, organized by the economists Colin Camerer and George Loewenstein. After a hiatus of several years this was followed by two meetings in 2001, one held by the Gruter Foundation for Law at their annual meeting in Squaw Valley. At that meeting the Gruter Foundation chose to focus its workshop on the intersection of neuroscience and economics, and invited several speakers active at the interface of these converging disciplines. The second meeting focused more directly on what would later become neuroeconomics, and was held at Princeton University. The meeting was organized by the neuroscientist Jonathan Cohen and the economist Christina Paxson, and is often seen as having been the inception of the present-day Society for Neuroeconomics. At this meeting, economists and neuroscientists met to explicitly discuss the growing convergence of these fields and to debate the value of such a convergence. There was, however, no consensus at the meeting that the growing convergence was desirable.

Nonetheless, the Princeton meeting generated significant momentum, and in 2003 a small invitation-only meeting that included nearly all of the active researchers in the emerging area was held on Martha's Vineyard, organized by Greg Berns of Emory University. This three-day meeting marked a clear turning point at which a group of economists, psychologists, and neurobiologists began to identify themselves as neuroeconomists and to explicitly shape the convergence between the fields. This led to an open registration meeting the following year at Kiawah Island, organized by Baylor College of Medicine's Read Montague. At this meeting a decision was made, by essentially all the central figures in the emerging discipline, to form a society and to turn this recurring meeting into an annual event that would serve as a focal point for neuroeconomics internationally. At the meeting, Paul Glimcher was elected President of the Society. The Society then held its first formal meeting in 2005 at Kiawah Island.

Against this backdrop of meetings, a series of critical papers and books was emerging that did even more to shape these interactions between scholars in the several disciplines, and to communicate the goals of the emerging neuroeconomic community to the larger neurobiological and economic communities. Probably the first neurobiological paper to rest explicitly on a normative economic theory was Peter Shizgal and Kent Conover's 1996 review, "On the neural computation of utility," in *Current Directions in Psychological Science*. This was followed the next year by a related paper published by Shizgal in *Current Opinion in Neurobiology* entitled "Neural basis of utility estimation." The reason that these papers can be viewed as the first in neuroeconomics is because they attempt to describe the neurobiological substrate of a behavioral choice using a form of normative choice theory derived from economics. In these papers, Shizgal analyzed the results of studies of intracranial self-stimulation in rats using a type of utility theory related loosely to the standard expected utility theory of von Neumann and Morgenstern. The papers argue that the choices an animal makes regarding whether or not to work for electrical stimulation of the medial forebrain bundle can be construed as an effort to maximize the animal's instant-to-instant utility. In this analysis, then, changes in the desirability of brain-stimulation reward as a function of stimulation frequency should be formally interpreted as changes in the utility of stimulus train. Unlike in standard theories of utility, however, Shizgal and Conover proposed that the expected utility of an action is perceived by the animal as the expected utility of that action divided by the sum of the expected utilities of all available actions. This particular formulation has its root in the work of the psychologist Richard Herrnstein, who proposed that many choices reflect this normalization with regard to the value of other alternatives – a phenomenon he referred to as *the matching law*. (For more about the matching law, see Chapter 30).

In fact, this equation had been introduced to self-stimulation studies five years earlier by Shizgal's mentor, C. Randy Gallistel. In the early 1990s, Gallistel had used Herrnstein's work to inspire quantitative choice-based experiments and analyses of intracranial self-stimulation (see Gallistel, 1994). Shizgal's extension of this work is critical in the history of neuroeconomics, because he moved away from the largely descriptive models of Herrnstein towards the normative models of economics. What Shizgal's work did not do, however, was fully incorporate the standard economic model, but rather a more normative version of Herrnstein's approach.

In 1999 this set of papers was followed by a paper by Platt and Glimcher (another student of Gallistel's) in *Nature* that argued quite explicitly for a normative utility-based analysis of choice behavior in monkeys (Platt and Glimcher, 1999). As they put it in that paper:

> Neurobiologists have begun to focus increasingly on the study of sensory-motor processing, but many of the models used to describe these processes remain rooted in the classic reflex … Here we describe a formal economic-mathematical approach for the physiological study of the sensory-motor process, or decision making.

At an experimental level, the paper goes on to demonstrate that the activity of single neurons in the posterior parietal cortex is a lawful function of both the probability and the magnitude of expected rewards. This was significant, because standard expected utility theory predicates choice on lawful functions of these same two variables. The paper, however, makes a critical mis-step in its examination of actual choice behavior. The authors go on to examine a matching-law type behavior which they interpret in terms of normative expected utility theory. This is problematic, because there is no normative standard for the analysis of matching-law behaviors. Indeed, in the example they present in the paper it cannot be proved that the behavior is predicted by their normative model; if anything, the data seem to suggest that the animals' behave sub-optimally. The result is a mixing of normative and non-normative approaches that characterized the early neurobiological work with economic approaches.

At the same time that this paper appeared in print, the behavioral economists Colin Camerer, George Lowenstein, and Drazen Prelec began circulating a manuscript in economic circles by the name of *Grey Matters*. In this manuscript the authors also argued for a neuroeconomic approach, but this time from a behavioral economic perspective. What these three economists argued was that the failures of traditional axiomatic approaches likely reflected neurobiological constraints on the algorithmic processes responsible for decision making. Neurobiological approaches to the study of decision, they argued, might reveal and define these constraints which cause deviations in behavior from normative theory.

What was striking about this argument, in economic circles, was that it proposed an algorithmic analysis of the physical mechanism of choice – a possibility that had been explicitly taboo until that time. Prior to the 1990s it had been a completely ubiquitous view in economic circles that models of behavior, like expected utility theory, were "as if" models – the model was to be interpreted "as if" utility were represented internally by the chooser. However, as Samuelson had argued half a century earlier, it was irrelevant whether this was actually the case because the models sought to link options to choices *not* to make assertions about the mechanisms by which that process was accomplished. Camerer and colleagues argued against this view, suggesting that deviations from normative theory should be embraced as clues to the underlying neurobiological basis of choice. In a real sense, then, these economists turned to neurobiology for exactly the opposite reason that the neurobiologists had turned to economics. They embraced neuroscience as a principled alternative to normative theory.

At this point, there was a rush by several research groups to perform an explicitly economic experiment that would mate these two disciplines in human choosers. Two groups succeeded in this quest in 2001. The first of these papers appeared in the journal *Neuron*, and reflected a collaboration between the functional magnetic resonance imaging pioneer Hans Breiter, Shizgal, and Kahneman (who would win the Nobel Prize in Economic Sciences for his contribution to behavioral economics the following year). This paper (Breiter *et al.*, 2001) was based on Kahneman and Tversky's *prospect theory*, a non-normative form of expected utility theory that guided much research in judgment and decision-making laboratories throughout the world (a theory described in detail in Chapter 11). In the paper, Breiter and colleagues manipulated the perceived desirability of a particular lottery outcome (in this case, winning zero dollars) by changing the values of two other possible lottery outcomes. When winning zero dollars is the worst of three possible outcomes, Kahneman and Tversky's prospect theory predicts that subjects should view it negatively; however, when it is the best of the three outcomes, then subjects should view it more positively. The scanning experiment revealed that brain activation in the ventral striatum matched these predicted subjective valuations.

The other paper published that year reflected a collaboration between the more neoclassically oriented economist Kevin McCabe, his colleague Vernon Smith (who would share the Nobel Prize with Kahneman the following year for his contributions to experimental economics), the econometrician Daniel Houser, and a team that included a psychologist and a biomedical engineer. Their paper, which appeared in the *Proceedings of the National Academy of Sciences of the United States of America* (McCabe *et al.*, 2001) examined behavior and neural activation while subjects engaged in a strategic game. This also represented the first use of game theory, an economic tool for the study of social decision making, in a neurobiological experiment. In this paper, subjects played a trust game either against an anonymous human opponent or against a computer, the details of which are reviewed in Chapter 5 of this volume. Their neurobiological data revealed that in some subjects the medial prefrontal cortex is differentially active under some of the conditions they examined, becoming more active when subjects play a cooperative strategy that deviates from the standard normative prediction of play in that game. From these data, the authors hypothesized that this non-normative pattern of cooperation has its origin in circuits of the prefrontal cortex.

The following year, many of these emerging trends were reviewed in an important special Society for Neuroscience conference issue of the journal *Neuron* (Volume 36, Issue 2) edited by Jonathan Cohen and Kenneth Blum entitled *Reward and Decision*. As these editors wrote in the introduction to that issue:

> Within neuroscience, for example, we are awash with data that in many cases lack a coherent theoretical understanding (a quick trip to the poster floor of the Society for Neurosciences meeting can be convincing on this point). Conversely, in economics, it has become abundantly evident that the pristine assumptions of the "standard economic model" – that individuals operate as optimal decision makers in maximizing utility – are in direct violation of even the most basic facts about human behavior.

In that issue, although all of the articles are by neurobiologists, particular attention is drawn to normative theories of decision. Of especial interest are articles by Montague and Berns (2002), Schultz (2002), Dayan and Balleine (2002), Gold and Shadlen (2002), and Glimcher (2002), which all point towards the interaction of normative models and neurobiology. Interestingly, the issue draws attention to the ongoing debate regarding the role of the neurotransmitter dopamine in reward processing, and draws upon previous work that had identified normative or near-normative models of learning that posit a role for dopamine. (This is a subject of tremendous importance to neuroeconomists today, and forms the focus of the third section of this volume.) What followed was a literal flood of decision-making studies in the neuroscientific literature, many of which relied on normative economic theory. Figure 1.1 documents this flood, plotting the number of papers published from 1990 to 2006 that list both "brain" and "decision making" as keywords.

At the end of this initial period, a set of summary reviews began to emerge that served as manifestos for the emerging neuroeconomic discipline. In 2003 Glimcher published a book, directed primarily at neuroscientists, that reviewed the history of neuroscience and argued that this history was striking in its lack of normative models for higher cognitive function (Glimcher, 2003). Glimcher proposed that economics could serve as the source for this much needed normative theory. Shortly thereafter the Camerer, Loewenstein, and Prelec paper was published under the title "Neuroeconomics" (Camerer *et al.*, 2005); this also served as a manifesto, but from the economic side.

Within the economic community a role similar to that of the *Neuron* special issue was played by a special issue on neuroeconomics presented by the journal *Games and Economic Behavior* (Volume 52, Issue 2) and edited by the economist Aldo Rustichini, which appeared shortly after this in 2005. Within the economic community this issue was hugely influential and served, to a large degree, to define neuroeconomics. The issue included articles by several economists and neuroscientists, including scholars ranging from Gallistel (2005) to Smith (Houser *et al.*, 2005).

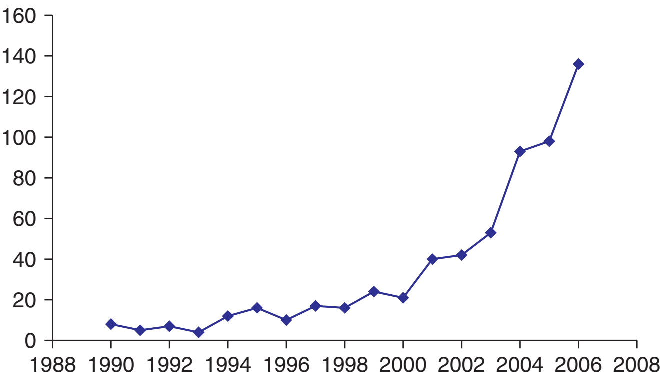


FIGURE 1.1 The increase in numbers of papers on decision-making studies in the neuroscientific literature, 1990–2006

Another major advance was presented in 2005, this one by Michael Kosfeld and his colleagues in Ernst Fehr's research group at the University of Zurich (Kosfeld *et al.*, 2005). This paper was important because it was the first demonstration of a neuropharmacological manipulation that alters behavior in a manner that can be interpreted with regard to normative theory. In the paper, subjects were asked to play a trust game much like the one examined by McCabe and colleagues. Fehr's critical manipulation was to increase brain levels of the neuropeptide oxytocin (by an intranasal application of the compound) before the players made their decision. What Kosfeld and colleagues found was that the investors with oxytocin sent more money to the trustees in the trust game than investors who received placebo. This increase in trusting behavior occurred despite the fact that investors' beliefs about the trustees' back-transfers remained unchanged. In contrast, oxytocin did not affect the trustees' behavior – i.e., trustees' back-transfers remained unchanged – ruling out the possibility that the neuropeptide just increases reciprocity or generosity. However, oxytocin did not cause an unspecific increase in the willingness to take risks, because in a control experiment – a pure risk game – the investors with oxytocin did not behave differently from the subjects with placebo. What was most interesting about this study from a neuroeconomic point of view was the demonstration that the administration of this endogenously produced neuropeptide altered a complex choice behavior of subjects in a very specific way – it neither affected the trustees' behavior nor did it affect the investors' general willingness to take risks, it only increased the investors' risk preference if the risk was constituted by the interaction with another human partner – suggesting a neurobiological basis for a difference between preferences for social and non-social risks.

The rise of neuroeconomics has been strongly associated with the rapid development of non-invasive neuroimaging techniques for human research and single-cell recordings in non-human primates. One limitation of these technologies is that they produce largely correlative measures of brain activity, making it difficult to examine the causal role of specific brain activations for choice behavior. This limitation can, however, be overcome with non-invasive methods of brain stimulation, such as transcranial magnetic stimulation (TMS) and transcranial direct current stimulation (tDCS), which enable researchers selectively to modify the neural processing associated with choice behavior. A recent study by Knoch *et al.* (2006) provides a demonstration of the additional neuroeconomic insights generated with these methods. Previous fMRI results (Sanfey *et al.*, 2003) had shown that the right *and* the left dorsolateral prefrontal cortex (DLPFC) are activated when subjects decide about the acceptance or rejection of unfair bargaining offers in the ultimatum game (for a description of this bargaining game, see Chapter 5). This finding raises many points, such as whether both hemispheres are causally involved in the choice process. Likewise, is DLPFC affecting judgments about the fairness of bargaining offers, or is it specifically involved in the implementation of fairness concerns? Knoch and colleagues disrupted the right and the left DLPFC with TMS and found that the disruption of both PFC areas left more abstract judgements of fairness fully intact (relative to a placebo stimulation), while the disruption of the right (but not the left) DLPFC resulted in a large increase in the acceptance of unfair offers. From a neuroeconomic viewpoint it is important to know the dissociations between judgment and choice, because choice typically implies that the decision maker must bear costs and benefits, while judgment alone is not yet associated with the bearing of costs and benefits. More generally, non-invasive brain stimulation techniques are likely to play an important role in future neuroeconomic studies because they provide causal knowledge and, in combination with imaging tools, make it possible to isolate whole decision networks that are causally involved in the generation of choices.

SUMMARY

Despite these impressive accomplishments, neuroeconomics is at best a decade old and has yet to demonstrate a critical role in neuroscience, psychology, or economics. Indeed, scholars within neuroeconomics are still debating whether neuroscientific data will provide theory for economists or whether economic theory will provide structure for neuroscience. We hope that both goals will be accomplished, but the exact form of this contribution is not yet clear. However, there are also skeptical voices, and the Pareto (1897) and Friedman arguments that economics is only about choices still lives in the form of fundamentalist critique. Gul and Pesendorfer (2008), for example, have argued that neuroscientific data and neuroscientific theories should, in principle, be unwelcome in economics.

The chapters that follow should allow readers to draw their own conclusions regarding this growing and dynamic field. Each of the major threads of contemporary research is reviewed in these pages. Although it is far too soon for there to be consensus in this community, the field today is small enough that a single volume can provide a comprehensive review. We therefore invite you, the readers, to estimate for yourselves the future directions that will yield greatest profit.

References

Allais, M. (1953). Le comportment de l'homme rationnel devant le risqué: Critique des postulates et axioms de l'ecole americaine. *Econometrica* 21, 503–546.

Bandettini, P.A., Wong, E.C., Hinks, R.S. *et al*. (1992). Time course EPI of human brain function during task activation. *Magn. Res. Med.* 25, 390–397.

Bechara, A., Damasio, H., Tranel, D., and Damasio, A. (1994). Deciding advantageously before knowing the advantageous strategy. *Science* 28, 1293–1295.

Breiter, H.C., Aharon, I., Kahneman, D. *et al*. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30, 619–639.

Bruni, L. and Sugden, R. (2007). The road not taken: how psychology was removed from economics, and how it might be brought back. *Economic J.* 117, 146–173.

Busino, G. (1964). Note bibliographique sur le Cours. In: V. Pareto (ed.), *Epistolario*. Rome: Accademia Nazionale dei Lincei, pp. 1165–1172.

Camerer, C., Loewenstein, G., and Prelec, D. (2005). Neuroeconomics: how neuroscience can inform economics. *J. Econ. Lit.* 43, 9–64.

Colander, D. (2007). Retrospectives: Edgeworth's hedonimeter and the quest to measure utility. *J. Econ. Persp.* 21, 215–225.

Dayan, P. and Balleine, B.W. (2002). Reward, motivation, and reinforcement learning. *Neuron* 36(2), 285–298.

Ellsberg, D. (1961). Risk, ambiguity and the savage axioms. *Q. J. Econ.* 75, 643–669.

Ferrier, D. (1878). *The Localization of Cerebral Disease*. New York, NY: G.P. Putnam and Sons.

Gallistel, C.R. (1994). Foraging for brain stimulation: toward a neurobiology of computation. *Cognition* 50, 151–170.

Gallistel, C.R. (2005). Deconstructing the law of effect. *Games Econ. Behav.* 52, 410–423.

Gigerenzer, G., Todd, P.M., and the ABC Research Group. (2000). *Simple Heuristics that Make Us Smart*. New York, NY: Oxford University Press.

Gilovich, T., Griffin, D., and Kahneman, D. (2002). *Heuristics and Biases: The Psychology of Intuitive Judgment*. New York, NY: Cambridge University Press.

Glimcher, P. (2002). Decisions, decisions, decisions: choosing a biological science of dhoice. *Neuron* 36, 323–332.

Glimcher, P. (2003). *Decisions, Uncertainty and the Brain: The Science of Neuroeconomics*. Cambridge, MA: MIT Press.

Glimcher, P.W. and Sparks, D.L. (1992). Movement selection in advance of action in the superior colliculus. *Nature* 355, 542–545.

Gold, J. and Shadlen, M. (2002). Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* 36, 299–308.

Green, D.M. and Swets, J.A. (1966). *Signal Detection Theory and Psychophysics*. New York, NY: Wiley.

Gul, F. and Pesendorfer, W. (2008). The case for mindless economics. In: A. Caplin and A. Schotter (eds), *The Foundations of Positive and Normative Economics: A Handbook*. Oxford: Oxford University Press, forthcoming.

Houser, D., Bechara, A., Keane, M. *et al*. (2005). Identifying individual differences: an algorithm with application to Phineas Gage. *Games Econ. Behav.* 52, 373–385.

Houthakker, H.S. (1950). Revealed preference and the utility function. *Economic*s 17, 159–174.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291.

Knoch, D., Pascual-Leone, A., Meyer, K. *et al*. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314, 829–832.

Kosfeld, M., Heinrichs, M., Zak, P.J. *et al*. (2005). Oxytocin increases trust in humans. *Nature* 435, 673–676.

Kwong, K.K., Bellieveau, J.W., Chesler, D.A. *et al*. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc. Natl Acad. Sci. USA* 89, 5675–5679.

Macmillan, M. (2002). *An Odd Kind of Fame: Stories of Phineas Gage*. Cambridge, MA: MIT Press.

McCabe, K., Houser, D., Ryan, L. *et al*. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl Acad. Sci. USA* 98, 11832–11835.

Montague, P.R. and Berns, G.S. (2002). Neural economics and the biological substrates of valuation. *Neuron* 36, 265–284.

Newsome, W.T., Britten, K.H., and Movshon, J.A. (1989). Neuronal correlates of a perceptual decision. *Nature* 341, 52–54.

Ogawa, S., Tank, D.W., Menon, R. *et al*. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proc. Natl Acad. Sci. USA* 89, 5951–5955.

Platt, M.L. and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.

Samuelson, P.A. (1938). A note on the pure theory of consumer behavior. *Economia* 1, 61–71.

Sanfey, A.G., Rilling, J.K., Aronson, J.A. *et al*. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science* 300, 1673–1675.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron* 36, 241–263.

Shizgal, P. (1997). Neural basis of utility estimation. *Curr. Opin. Neurobiol.* 7, 198–208.

Shizgal, P. and Conover, K. (1996). On the neural computation of utility. *Curr. Direct. Psycholog. Sci.* 5, 37–43.

Smith, V. (1976). Experimental economics: induced value theory. *Am. Econ. Rev.* 66, 274–279.

Sparks, D.L. (1999). Conceptual issues related to the role of the superior colliculus in the control of gaze. *Curr. Opin. Neurobiol.* 9, 698–707.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

PART 1

NEOCLASSICAL ECONOMIC APPROACHES TO THE BRAIN

CHAPTER

2

Introduction: Experimental Economics and Neuroeconomics

Vernon L. Smith

OUTLINE

INTRODUCTION

There are three interdependent orders of brain/mind decision making that I believe are essential to our understanding of the human career: first, the internal order of the mind, the *forte* of neuroscience from its inception; second, the external order of socio-economic exchange, which constitutes the reciprocity and sharing norms that characterize human sociality as a cross-cultural universal; and third, the extended order of cooperation through market institutions and technology. This is the foundation of wealth creation through specialization whose ancient emergence is manifest on a global scale.

The social brain seems to have evolved adaptive mechanisms for each of these tasks, which involve experience, memory, perception, and personal tacit knowledge, or "can do" operating skill. This theme was prominent in the reflections and observations on human sociality (sentiments, empathy) and market order, respectively, in the two works of Adam Smith (1759, 1776) and subsequently in Darwin's celebrated biological perspective on psychology as arising from "the acquirement of each mental power and capacity by gradation" in evolutionary time (Darwin, 1859: 458). That acquirement process, I believe, significantly relates to human decision-making capacity within cultural constraints – the norms and rules of engagement of the local social order, or "order without law" (Ellickson, 1991) – and the more formal rules of law that constrain decision in the extended order of market institutions.

Less obvious, but perhaps even more important to an understanding of the human enterprise, has been the process of cultural change in the norms and rules that constrain decision, and in the evolution of institutions that govern the market order.

Experimental economics has been driven by the power of using controlled experiments, both in the laboratory and in the field, to illuminate the study of these three orders of human interactive decision. Neuroeconomics adds new brain-imaging and emotion-recording technologies for extending and deepening these investigations. Consequently, it offers much promise for changing the way we think about and research brain function, decision, and human sociality. It will surely chart entirely new if unpredictable directions once the pioneers get beyond trying to find better answers to the questions they inherited from the past. Neuroeconomic achievement is most likely to be recognized for its ability to bring a new perspective and understanding to the examination of important economic questions that have been intractable or beyond the reach of traditional economics. Initially, new tools tend naturally to be applied to the old questions; however, their ultimate importance emerges when the tools change how people think about their subject matter, enable progress on previously unimaginable new questions, and lead to answers that would not have been feasible before the innovation. Neureconomics will be known by its deeds, and no one can foresee the substance of those deeds.

Neuroeconomics has this enormous nonstandard potential, but it is far too soon to judge how effective it will be in creating new pathways of comprehension.

In this spirit, I propose in this short introduction to probe those areas of economics that I believe are in most need of fresh new empirical investigation and understanding; areas where our interpretation of experimental results depends on assumptions that are difficult to test, but which may yield to neuroeconomic observations.

THE INTERNAL ORDER: REWARDS AND THE BRAIN

Neuroscience has been particularly useful in deepening our perspectives on questions related to motivation and the theory of choice – for example, does the brain encode comparisons on a relative or an absolute scale? Animal studies of choice show that they respond to pair-wise comparisons of differential rewards. It is now established that orbital frontal cortex neuron activity in monkeys enables them to discriminate between rewards that are directly related to the animals' relative (as distinct from absolute) preference among food item such a cereal, apples, and raisins (in order of increasing preference) (Tremblay and Schultz, 1999). Thus, if A is preferred to B is preferred to C, then neuronal activity is greater for A than for B when the subject is comparing A and B, and similarly for B and C when comparing B and C. But the amplitude intensity associated with B is much greater when compared to C than when it is compared to A, which is contrary to what might be expected if A, B, and C were encoded on a fixed scale of values rather than a relative scale (Tremblay and Schultz, 1999: 706).

Choice behavior, however, may relate to perception (e.g., orbital frontal response) differently from how it relates to individual utility value in problem-solving (e.g., parietal response). Glimcher (2003: 313–317) reports studies in which a monkey chooses between two options ("work or shirk") in a Nash game against a computer. The choice *behavior* of the monkey tracks changes in the Nash equilibrium prediction in response to changes in the outcome payoffs. However, neuron (LIP) firing in the parietal cortex does *not* track the changing equilibrium values, but remains steady at the relative (unchanging) realized *expected* payoffs such that the decision maker is indifferent between the options available – i.e., the expected payoffs are the same in the comparison. These results are consistent with the hypothesis that the brain computes and maintains equilibrium while behavior responds to changes in the payoffs.

These studies appear to have parallel significance for humans. In prospect theory, the evaluation of a gamble depends not on the total asset position but marginally on the opportunity cost, gain or loss, relative to a person's baseline current asset position. Moreover, as noted by Adam Smith (1759; 1982: 213), the effect of a loss looms larger than the effect of the gain – a robust phenomenon empirically established by Kahneman and Tversky (1979). Similarly, Mellers *et al*. (1999) found that the emotional response to a gamble's outcome depends on the perceived value and likelihood of the outcome, but also on the foregone outcome. It feels better (less bad) to receive \$0 from a gamble when you forgo +\$10 than when you forgo +\$90. Opportunity cost comparisons for decision are supported by our emotional circuitry, and that support is commonly below our conscious awareness.

The human brain acquired its reward reinforcement system for food, drink, ornaments, and other items of cultural value long before money was discovered as a mechanism for facilitating exchange. Consequently, our brains appear to have adapted to money, as an object of value, or "pleasure," by simply treating it like another "commodity," latching on to the older receptors and reinforcement circuitry (Thut *et al*., 1997; Schultz, 2000, 2002). To the brain, money is like food, ornaments, and recreational drugs, and only indirectly signals utility derived from its use. However, this interpretation is conditional on an external context in which the exchange value of money is stable. We need to learn how the brain adapts when the context changes: How do our brains monitor and intervene to modify this reinforcement when money is inflated by monetary authorities, sometimes eroding to worthlessness?

Money is a social contrivance, and in its exchange value to the individual we see human sociality at work; we accept fiat money only so long as we have confidence that others will accept it from us. We also see sociality at work in individual decision based on observing and learning from the experience of others. Hence, individual decision modeled as a game against nature does not imply social isolation.

THE SOCIAL ORDER

These considerations leave unanswered questions regarding how to interpret behavior in social interactions as observed in a variety of two-person experimental games.

Thus, in extensive-form trust games played anonymously only once, people cooperate more than is predicted by game theory based on the hypotheses that people choose according to dominance criteria, perceive single-play games as events isolated from all other social interactions, and always apply backward induction to analyze decisions. Is cooperation motivated by altruism (social preferences), by the personal reward that emanates from relationship-building (goodwill) in exchange, or by failure to backward induct? (See Chapter 15 of this volume; McCabe and Smith, 2001; Johnson *et al.*, 2002.) Many experiments have sought to explore or reduce this confounding (see Smith, 2008: 237–244, 257–264, 275–280, for summaries and references).

Repeated games are modeled by assuming that individual (i) with current utility (ui) chooses strategy (si) in a stage game to maximize (1 − d)ui(s) + d Vi(H(s)), where s = (s1, …si, …sn), d is a discount factor, n is the number of players, H is the history of play, and d Vi(H) is i's endogenous subjective discounted value of continuation (Sobel, 2005). Hence, the continuation value perceived by i may easily make it in her interest to forgo high current utility from domination because it reduces the value she achieves in the future. An open question is how individuals perceive Vi. We ordinarily think that our procedures for implementing single play should yield Vi = 0, and the choice is the dominant immediate payoff from si. But is this assumption defensible?

In a trust game (with n = 2), a cooperative response by the second player has been discovered to depend on the opportunity cost to the first player of choosing to offer the prospect of cooperation. Second movers defect twice as often when they see that the first player has no option but to pass to the second versus seeing the sure-thing payoff given up by the first in order to enable cooperation and a greater reward for both (McCabe *et al.*, 2003). Thus, defection in favor of payoff dominance is much reduced when the circumstances suggest intentional action at a cost to the first player in facilitating cooperative gains from the exchange. Moreover, fMRI data confirm that circuitry for detecting intentions is indeed activated in trust games (McCabe *et al.*, 2001). Knowing that you did something costly for me may increase my unconscious motivation to reciprocate your action, to implicitly recognize a relationship, as in day-to-day socializing. Hence, forgoing u(si) is part of H in a sequential move single-play game, and the players need not be oblivious to this "history" if they share common cultural experiences.

Many other experiments report results consistent with relationship-building in single-play stage games, suggesting that we have failed to implement the key conceptual discontinuity in game theory between single and repeat play. For example:

1. In dictator games, altruistic behavior substantially increases when people are informed that a second undefined task will follow the first. This discovery shows how strongly people can be oriented unconsciously to modify their behavior if there is any inadvertent futurity in the context (Smith, 2008: 237–244).
2. Double-blind protocols affect behavior in single-play dictator and trust games, establishing that people are more cooperative when third parties can know their decisions – a condition that we would expect to be important in reputation-building only when there is repeat interaction.
3. People are more cooperative when the "equivalent" stage game is played in extensive rather than abstract strategic form. The latter is rare in everyday life, but is very convenient for proving theorems. The extensive form triggers an increase in cooperative behavior consistent with the discussion above, although own and other payoffs are identical in the comparisons (see Smith, 2008: 264–267, 274–275, for a summary and references).

As experimentalists, we have all become comfortable with our well-practiced tool kits for implementing and rewarding subjects in single-play games. But are experimental results affected by an "other people's money" (OPM) problem when the experimenter gift-endows the subjects up front? Cherry *et al.* (2002) show that dictator-game altruism all but disappears (97% of the subjects give nothing) under double-blind conditions when subjects are first required to earn their own endowments. Oxoby and Spraggon (2008) vary these protocols and report that 100% of their dictators give nothing. These results raise fundamental questions concerning the effect of OPM on observed social behavior in the laboratory.

If cooperation derives from social preferences, the neuroeconomic question is: how is the encoding of payoff to self and to other affected by the circumstances under which the resources are acquired? Yes, we know that the "same" general OFC area is activated by own and other rewards (see Chapters 15 and 20 in this volume), but how do those activations network with other brain areas and respond to differential sources of reward monies?

Hence there is the prospect that neuroeconomics can help to disentangle confounding interpretations of behavior in trust and other two-person games. But in this task we need continually to re-examine and challenge our implicit suppositions to avoid painting ourselves into a confirmatory corner. In summary:

- How are brain computation processes affected by *who* provides the money or *how* people acquired the stakes – the OPM problem?
- When people cooperate, which game theoretic hypothesis do we reject: payoff dominance independent of circumstances, or our procedures for implementing the abstract concept of a single-play game? If the former, we modify preferences; if the latter, we have to rethink the sharp distinction in theory between single-play games and goodwill-building in repeat interaction. Either or both may be manifestations of the social brain.
- How does relationship building differ between people who do and those who do not naturally apply backward induction to the analysis of decisions?
- What accounts for the behavioral non-equivalence between the extensive and normal form of a game?

In the absence of a deeper examination of these questions, we cannot (1) distinguish between exchange and preference interpretations of human sociality; (2) understand why context is so important in determining cooperative behavior; (3) understand how cooperation is affected by repeat play across the same or different games, under different subject matching protocols.

THE MARKET ORDER

Hundreds of market experiments have demonstrated the remarkable ability of unsophisticated subjects to discover equilibrium states, and to track exogenous changes in these states over time in repeat interaction with small numbers under private information. Yet the mental processes that explicate this skill are unavailable to those who demonstrate the ability, are inadequately modeled, and understood only as empirical phenomena. For example it is well documented that individual incentives and the rules of market institutions matter in determining the speed and completeness of equilibrium convergence, but it is unknown how the brain's decision algorithms create these dynamic outcomes.

Can neuroeconomics contribute to an understanding of how uncomprehending individual brains connect with each other through rules to solve the market equilibrium problem and in this process create wealth through specialization? Lab experiments, betting markets, many information markets, and some futures markets demonstrate how effective people are at efficiently aggregating dispersed information.

We also have the puzzle that, under both the explicit property right rules of the market order and the mutual consent reciprocity norms of the social order, the individual must give in order to receive in exchange. However, this insight is not part of the individual's perception. The individual, who perceives social as well as self-betterment through cooperation in personal exchange with others, does not naturally see the same mechanism operating in impersonal market settings. Yet individuals in concert with others create both the norms of personal exchange and the institutions of market exchange.

In closing, it is evident that there is mystery aplenty across the spectrum of individual, social, and market decision to challenge the discovery techniques of neuroeconimcs. However, to meet that challenge I believe we must be open to the exploration of assumptions that underpin theory and experiment. I am optimistic about this prospect and how it may contribute to future understanding.

References

Cherry, T., Frykblom, P., and Shogren, J. (2002). Hardnose the dictator. *Am. Econ. Rev.* 92, 1218–1221.

Darwin, C. (1859) (1979 edn). *The Origin of Species*. Darby: Arden Library.

Ellickson, R.C. (1991). *Order Without Law: How Neighbors Settle Disputes*. Cambridge, MA: Harvard University Press.

Glimcher, P. (2003). *Decisions, Uncertainty and the Brain*. Cambridge, MA: MIT Press.

Johnson, E., Camerer, C., Sen, S., and Rymon, T. (2002). Detecting failures of backward induction: monitoring information search in sequential bargaining environments. *J. Econ. Theory* 104, 16–47.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291.

McCabe, K., Houser, D., Ryan, L. *et al*. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl Acad. Sci. USA* 98, 11832–11835.

McCabe, K., Rigdon, M., and Smith, V. (2003). Positive reciprocity and intentions in trust games. *J. Econ. Behav. Org.* 52(2), 267–275.

McCabe, K. and Smith, V. (2001). Goodwill accounting and the process of exchange. In: G. Gigerenzer and R. Selten (eds), *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press, pp. 319–340.

Mellers, B., Schwartz, A., and Ritor, I. (1999). Emotion-based choice. *J. Exp. Psychol. Gen.* 128, 1–14.

Oxoby, R. and Spraggon, J. (2008). Mine and yours: property rights in dictator games. *J. Econ. Behav. Org.* 65(3–4), 703–713.

Schultz, W. (2000). Multiple reward signals in the brain. *Nature Rev. Neurosci.* 1, 199–207.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron* 36, 241–263.

Smith, A. (1759) (1982 edn). *The Theory of Moral Sentiments* (D. Raaphaet and A. Mactie, eds). Indianapolis: Liberty Fund.

Smith, A. (1776) (1981 edn). *An Enquiry into the Nature and Causes of the Wealth of Nations*, Vol. 1 (R. Campbell and A. Skinner eds). Indianapolis, IN: Liberty Fund.

Smith, V. (2008). *Rationality in Economics: Constructivist and Ecological Forms*. Cambridge: Cambridge University Press.

Sobel, J. (2005). Interdependent preferences and reciprocity. *J. Econ. Lit.* 42, 392–436.

Thut, G., Schultz, W., Roelcke, U. *et al*. (1997). Activation of the human brain by monetary reward. *NeuroReport* 8, 1225–1228.

Tremblay, L. and Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature* 389, 704–708.

CHAPTER

3

Axiomatic Neuroeconomics

Andrew Caplin and Mark Dean

OUTLINE

INTRODUCTION

Those of us who pursue neuroeconomic research do so in the belief that neurobiological and decision theoretic research will prove highly complementary. The hoped for complementarities rest in part on the fact that model-building and quantification are as highly valued within neuroscience as they are in economics. Yet methodological tensions remain. In particular, the "axiomatic" modeling methodology that dominates economic decision theory has not made many neuroscientific converts. We argue in this chapter that neuroeconomics will achieve its full potential when such methodological differences are resolved, and in particular that axioms can and should play a central role in the development of neuroeconomics.

The axiomatic approach to modeling is the bread and butter of decision theory within economics. In pursuing this approach, model-builders must state precisely how their theories restrict the behavior of the data they are interested in. To make such a statement, the model-builder must write down a complete list of necessary and sufficient conditions (or axioms) that his data must satisfy in order to be commensurate with his model. The classic example in decision theory (which we discuss more in following section) is the case of "utility maximization." While this had been the benchmark model of economic behavior almost since the inception of the field, it was left to Samuelson (1938) to ask the question: "Given that we do not observe 'utility', how can we test whether people are utility maximizers?" In other words: What are the observable characteristics of a utility maximizer? It turns out that the answer is the *Weak Axiom of Revealed Preference* (WARP), which effectively states that if someone chooses some option x over another y, he cannot later be observed choosing y over x. If (and only if) this rule is satisfied, then we can say that the person in question behaves as if choosing in order to maximize some fixed, underlying utility ordering. Although this condition may seem surprisingly weak, it is the *only* implication of utility maximization for choice, assuming that utility is not directly observed. Furthermore, it turns out that there are many cases in which it systematically fails (due, for example, to framing effects, *status quo* bias or "preference reversals"). In the wake of this pivotal insight, the axiomatic approach has been successfully used within economics to characterize and test other theories which, like utility maximization make use of "latent" variables (those which are not directly observable), sometimes called intervening variables.

It is our belief that axiomatic modeling techniques will prove to be as valuable to neuroeconomics as they are to economics. As with utility, most of the concepts studied in neuroeconomics are not subject to direct empirical identification, but can be defined only in relation to their implications for particular neurological data. Axioms are unique in the precision and discipline that they bring to modelling such latent forces, in that they capture *exactly* what they imply for a particular data set – no more and no less. Moreover, they capture the main characteristics of a model in a non-parametric way, thus removing the need for "spurious precision" in relating latent variables to observables – as well as the need for the many free parameters found in a typical neurobiological model. An axiomatic approach also fixes the meaning of latent variables by defining them relative to the observable variables of interest. This removes the need for auxiliary models, connecting these latent variables to some other observable in the outside world. In the third section of this chapter, we illustrate our case with the neurobiological/neuroeconomic question of whether or not dopamine encodes a "reward prediction error" (Caplin and Dean, 2008b; Caplin *et al.*, 2008a). We show the value of an axiomatic model in identifying the latent variables *rewards* and *beliefs* in terms of their impact on dopaminergic responses, just as revealed-preference theory identifies utility maximization relative to its impact on choice.

Note that we see the use of axiomatic methods not as an end in and of itself, but rather as a guide to drive experimentation in the most progressive possible directions. Not only do good axiomatic models immediately suggest experimental tests; they also lend themselves to a "nested" technique of modeling and experimentation, in which successively richer versions of the same model can be tested one step at a time. Ideally, this creates rapid feedback between model and experiment, as refinements are made in the face of experimental confirmation, and adjustments in the face of critical contrary evidence. This nested modeling technique results in a shared sense of the challenges that stand in the path of theoretical and empirical understanding.

One reason that this approach has proven so fruitful in economics is that our theories are very far from complete in their predictive power. There is little or no hope of constructing a simple theory that will adequately summarize all relevant phenomena; systematic errors are all but inevitable. The axiomatic method adds particular discipline to the process of sorting between such theories. In essence, the key to a successful axiomatic agenda involves maintaining a close connection between theoretical constructs and empirically observable phenomena.

Overall, axiomatic modeling techniques strike us as an intensely practical weapon in the neuroscientific arsenal. We are driven to them by a desire to find good testing protocols for neuroeconomic models, rather than by a slavish devotion to mathematical purity. In addition to operationalizing intuitions, axioms allow the capture of important ideas in a non-parametric way. This removes the need for overly specific instantiations, whose (all but inevitable) ultimate rejection leaves open the possibility that the intuitive essence of the model can be retained if only a better-fitting alternative can be found in the same model class. By boiling a model down to a list of necessary and sufficient conditions, axioms allow identification of definitive tests. With the implied focus on essentials and with extraneous parametric assumptions removed from the model, failure to satisfy the axioms implies unequivocally that the model has problems which go far deeper than a particular functional form or set of parameter values. The rest of this essay illustrates these points. In the following section, we discuss briefly the success that the axiomatic method has had within economics. We then discuss some of our own work in applying the same methodology to a neurobiological/neuroeconomic question: whether or not dopamine encodes a "reward prediction error." We conclude by outlining some next steps in the axiomatic agenda in neuroscience.

THE AXIOMATIC METHOD IN DECISION THEORY

Within decision theory, axiomatic methods have been instrumental to progress. It is our contention that neuroeconomic applications of this approach are highly promising, for exactly the same reasons that they have proven so fruitful in economics. In essence, the key to a successful axiomatic agenda involves maintaining a close connection between theoretical constructs and empirically observable phenomena. A quick review of doctrinal history highlights the possible relevance of these techniques for neuroeconomics.

In general, the starting point for an axiomatic theory in economics has been an area in which strong intuitions about the root causes of behavior are brought to play, and in which questions arise concerning how these intuitive causes are reflected in observables. This interplay between theory and data was evident from the first crucial appearance of axiomatic methods in economics: the revealed preference theory initiated by Paul Samuelson.

The debate which gave birth to the revealed-preference approach, and so axiomatic modeling within economics, goes back to the beginning of economic thought, and the question of what determines observed market prices. The notion of "use value," or the intrinsic value of a good, was central in early economics, with debates focusing on how this related to prices. The high price of diamonds, which seem to have low use value, relative to water, which is necessary for sustaining life, was seen as a source of great embarrassment for proponents of the idea that prices reflected subjective evaluations of the relative importance of commodities. Understanding of the connection between this early notion of "utility" and prices was revolutionized when marginal logic was introduced into economics in the late nineteenth century. It was argued that prices reflect marginal, not total, utilities (i.e. the incremental utility of owning an additional unit of a commodity), and that marginal utility fell as more of a commodity became available. Water is abundant, making marginal units of low value. However, if water were to be really scarce, its market value would increase tremendously to reflect the corresponding increase in marginal utility. Thus, if water were as scarce as diamonds, it would be far more valuable.

There were two quite different responses to this theoretical breakthrough, one of which led to a long philosophical debate that has left little mark on the profession, and the other of which produced the most fundamental axiomatic model in choice theory. The philosophical response was produced by those who wanted to dive more fully into the sources and nature of utility, whether or not it really diminished at the margin, and what form of "hedonometer" could be used to measure it. It could be argued that the form of utility offered by diamonds is fundamentally different than that offered by water: diamonds may be of value in part because of their scarcity, while water is wanted for survival. One could further reflect philosophically on how well justified was each such source of utility, how it related to well-being, and why it might or might not decrease at the margin. The alternative, axiomatic response resulted when those of a logical bent strove to strip utility theory of inessential elements, beginning with Pareto's observation that the utility construct was so flexible that, the concept that it diminished at the margin was meaningless: the only legitimate comparisons, he argued, involve better than, worse than, and indifferent to – information that could be captured in an ordinal preference ranking[1]. This observation made the task of finding "the" measurable counterpart to utility seem inherently hopeless, and it was this that provoked Paul Samuelson to pose the fundamental question concerning revealed preference that lies at the heart of modern decision theory.

Samuelson noted that the information on preferences on which Pareto proposed building choice theory was no more subject to direct observation than were the utility functions that were being sought by his precursors: neither preferences or utilities are directly observable. In fact, the entire content of utility maximization theory seemed purely intuitive, and Samuelson remarked that there had been no thought given to how this intuitive concept would be expected to play out in observed choices. His advance was to pose the pivotal question precisely: if decision makers are making choices in order to maximize *some* utility function (which we cannot see), what rules do they have to obey in their behavior? If the theory of utility maximization had been shown to have no observable implications for choice data, Samuelson would have declared the concept vacuous.

In a methodological achievement of the first order, it was shown by Samuelson and others that utility maximization does indeed have such implied restrictions. These are identified precisely by the Weak Axiom of Revealed Preference. In the simplest of case, the axiom states essentially that if I see you choose some object x over another object y, I cannot in some other experiment see you choose y over x. If and only if this condition holds do you behave as if making choices in order to maximize some fixed utility function. The broader idea is clear. This *revealed preference* (Samuelson favored "revealed chosen") methodology calls for theory to be tied closely to observation: utility maximization is defined only in relation to the observable of interest – in this case, choice. There is no need for additional, auxiliary assumptions which tie utility to other observables (such as "amount of food" or "softness of pillow"). Furthermore, the approach gives insights into the limits of the concept of utility. As utility only represents choice, it is only defined in the sense that it represents an ordering over objects: it does not provide any cardinal information. In other words, any utility function which preserves the same ordering will represent choice just as well; we can take

---

[1] An "ordinal" relation is one which includes only information on the ranking of different alternatives, as opposed to a "cardinal" relation which contains information about *how much* better one alternative is than another.

all utility values and double them, add 5 to them, or take logs of them, and they will all represent the same choice information. It is for this reason that the concept of utility diminishing at the margin is meaningless: for any utility function which shows diminishing marginal utility we can find another one with increasing marginal utility which represents choice just as well[2].

To understand how best to apply the axiomatic methodology, note that Samuelson was looking to operationalize the concept of utility maximization, which has strong intuitive appeal. Having done so, the resulting research agenda is very progressive. The researcher is led to exploring empirical support for a particular restriction on choice data. Where this restriction is met, we can advance looking for specializations of the utility function. Where this restriction is not met, we are directed to look for the new factors that are at play that by definition cannot be covered by the theory of utility maximization. After 150 years of verbal jousting, revealed preference theory put to an end all discussion of the purview of standard utility theory and moreover suggested a progressive research program for moving beyond this theory in cases in which it is contradicted. Ironically, it has taken economists more than 60 years to follow up on this remarkable breakthrough and start to characterize choice behaviors associated with non-maximizing theories.

The area of economics in which the interplay between axiomatic theories and empirical findings has been most fruitful is that of decision making under uncertainty. The critical step in axiomatizing this set of choices was taken by von Neumann and Morgenstern (1944), who showed that a "natural" method of ranking lotteries[3] according to the expected value of a fixed reward function (obtained by multiplying the probability of obtaining each outcome with the reward associated with that outcome) rests on the highly intuitive substitution, or independence axiom. This states that if some lottery p is preferred to another lottery q, then the weighted average of p with a third lottery r must be preferred to the same weighting of q with r.

This theory naturally inspired specializations for particular applications as well as empirical criticisms. Among the former are the theory of risk aversion (Pratt, 1964), and asset pricing (Lucas, 1971), which now dominate financial theory. Among the latter are such behaviors as those uncovered by Allais (1953), Ellsberg (1961), Kahneman and Tversky (1973), and various forms of information-seeking or information-averse behavior. These have themselves inspired new models based on different underlying axiom sets. Examples include models of ambiguity aversion (Schmeidler, 1982; Gilboa and Schmeidler, 1989), disappointment aversion (Gul, 1991), rank-dependent expected utility (Quiggin, 1982), and preferences over the date of resolution of uncertainty (Kreps and Porteus, 1979).

The interaction between theory and experimentation has been harmonious due in large part to the intellectual discipline that the axiomatic methodology imposes. Theory and experimentation ideally advance in a harmonious manner, with neither getting too far ahead of the other. Moreover, as stressed recently by Gul and Pesendorfer (2008), axiomatic methods can be used to discipline the introduction of new psychological constructs, such as anxiety, self-control, and boundedly rational heuristics, into the economic canon. Rather than simply naming these latent variables in a model and exploring implications, the axiomatic method calls first for consideration of precisely how their inclusion impacts observations of some data set (albeit an idealized data set). If their inclusion does not after the range of predicted behaviors, they are not seen as "earning their keep." If they do increase the range of predictions, then questions can be posed concerning when and where such observations are particularly likely. Thus, the axiomatic method can be employed to ensure that any new latent variable adds new empirical predictions that had proven hard to rationalize in its absence[4].

AXIOMS AND NEUROECONOMICS: THE CASE OF DOPAMINE AND REWARD PREDICTION ERROR

Among the parent disciplines of neuroscience are physics, chemistry, biology, and psychology. Quantitative modeling abounds in the physical sciences, and this is mirrored in various areas of neuroscience, such as in the field of vision. Yet there

---

[2] What is meaningful is whether the rate at which a decision maker will trade one good off against another – the marginal rate of substitution – is increasing or decreasing.

[3] Note that economists conceptualize choice between risky alternatives as a choice between lotteries. Each lottery is identified with a probability distribution over possible final outcomes. Such a lottery may specify, for example, a 50% chance of ending up with $100 and a 50% chance of ending up with $50.

[4] The axiomatic method does not call for the abandonment of common sense. After all, we can provide many axiomatizations of the same behavior involving quite different latent variables, and an esthetic sense is used in selecting among such axiomatizations. Yet anyone who wishes formally to reject one equivalent axiomatization over another must identify a richer setting in which they have distinct behavioral implications.

remain many psychological constructs that have been imported into behavioral neuroscience which, while subject to powerful intuition, continue to elude quantification. These include motivation, cognitions, construal, salience, emotions, and hedonia.

An element shared by the disciplines out of which neuroscience has evolved is that axiomatic methods either have been entirely neglected, or are seen as having contributed little to scientific progress. In particular axiomatic methods have earned something of a bad name in psychological theory, in which their use has not been associated with a progressive interaction between theory and data. Within the physical sciences, the data are so rich and precise that axioms have typically been inessential to progress. However, we believe that neuroeconomics is characterized by the same combination of conditions that made the axiomatic method fruitful within economics. Intuition is best gained by working with concepts such as "reward," "expectations," "regret," and so on, but the exact relation of these concepts to observables needs to be made more precise. It is the axiomatic method that allows translation of these intuitive notions into *observable* implications in as clear and general a manner as possible.

We illustrate our case with respect to the neurotransmitter dopamine. The *reward prediction error* model (RPE) is the most well-developed model of dopaminergic function, and is based on such intuitive concepts as rewards and beliefs (i.e., expectations of the reward that is likely to be obtained in a particular circumstance). Yet, as in the case of utility theory, these are not directly observable. Commodities and events do not come with readily observable "reward" numbers attached. Neither are beliefs subject to direct external verification. Rather, both are latent variables whose existence and properties must be inferred from a theory fit to an experimental data set. The natural questions in terms of an axiomatic agenda are analogous to those posed in early revealed-preference theory: what restrictions does RPE model place on observations of dopamine activity. If there are no restrictions, then the theory is vacuous. If there are restrictions, are the resulting predictions verified? If so, is it possible to develop further specializations of the theory that are informative on various auxiliary hypotheses? If not, to what extent can these be overcome by introducing particular alternative theories of dopaminergic function? This is precisely the agenda that we have taken up (Caplin and Dean, 2008b; Caplin *et al.*, 2008a), and to which we now turn.

A sequence of early experiments initially led neuroscientists to the conclusion that dopamine played a crucial role in behavior by mediating "reward." Essentially, the idea was that dopamine converted experiences into a common scale of "reward" and that animals (including the human animal) made choices in order to maximize this reward (see, for example, Olds and Milner, 1954; Kiyatkin and Gratton, 1994; see also Gardner and David, 1999, for a review). The simple hypothesis of "dopamine as reward" was spectacularly disproved by a sequence of experiments highlighting the role of *beliefs* in modulating dopamine activity: whether or not dopamine responds to a particular reward depends on whether or not this reward was *expected*. This result was first shown by Schultz and colleagues (Schultz *et al.*, 1993; Mirenowicz and Schultz, 1994; Montague, Dayan, and Sejnowski, 1996). The latter study measured the activity of dopaminergic neurons in a thirsty monkey as it learned to associate a tone with the receipt of fruit juice a small amount of time later. Initially (i.e., before the animal had learned to associate the tone with the juice), dopamine neurons fired in response to the *juice* but not the *tone*. However, once the monkey had learned that the tone predicted the arrival of juice, then dopamine responded to the tone, but now did *not* respond to the juice. Moreover, once learning had taken place, if the tone was played but the monkey did not receive the juice, then there was a "pause" or drop in the background level of dopamine activity when the juice was expected.

These dramatic findings concerning the apparent role of information about rewards in mediating the release of dopamine led many neuroscientists to abandon the hedonic theory of dopamine in favor of the RPE hypothesis: that dopamine responds to the difference between how "rewarding" an event is and how rewarding it was expected to be[5]. One reason that this theory has generated so much interest is that a reward prediction error of this type is a key algorithmic component of reward prediction error models of learning: such a signal is used to update the value attached to different actions. This has led to the further hypothesis that dopamine forms part of a reinforcement learning system which drives behavior (see, for example, Schultz *et al.*, 1997).

The RPE hypothesis is clearly interesting to both neuroscientists and economists. For neuroscientists, it offers the possibility of understanding at a neuronal level a key algorithmic component of

---

[5] The above discussion makes it clear that reward is used in a somewhat unusual way. In fact, what dopamine is hypothesized to respond to is effectively unexpected changes in lifetime "reward:" dopamine responds to the bell not because the bell itself is rewarding, but because it indicates an increased probability of future reward. We will return to this issue in the following section.

the machinery that governs decision making. For economists, it offers the opportunity to obtain novel insight into the way beliefs are formed, as well as further develop our models of choice and learning. However, the RPE hypothesis is far from universally accepted within the neuroscience community. Others (e.g., Zink *et al.*, 2003) claim that dopamine responds to "salience," or how surprising a particular event is. Berridge and Robinson (1998) claim that dopamine encodes "incentive salience," which, while similar to RPE, differentiates between how much something is "wanted" and how much something is "liked." Alternatively, Redgrave and Gurney (2006) think that dopamine has nothing to do with reward processing, but instead plays a role in guiding attention. Developing successful tests of the RPE hypothesis which convince all schools is therefore a "neuroeconomic" project of first-order importance. Developing such tests is complicated by the fact that the RPE model hypothesizes that dopamine responds to the interaction of *two* latent (or unobservable) variables: reward and beliefs. Anyone designing a test of the RPE hypothesis must first come up with a solution to this quandary: how can we test whether dopamine responds to changes in things that we cannot directly measure?

The way that neuroscientists studying dopamine currently solve this latent variable problem is by adding to the original hypothesis further models which relate beliefs and rewards to observable features of the outside world. More specifically, "reward" is usually assumed to be linearly related to some "good thing," such as fruit juice for monkeys, or money for people. Beliefs are usually calibrated using a reward prediction error model. Using this method, for any given experiment, a time series of "reward prediction error" can be generated, which can in turn be correlated with brain activity. This is the approach taken in the majority of studies of dopamine and RPE in monkeys and humans (see, for example, Montague and Berns, 2002; O'Doherty *et al.*, 2003, 2004; Bayer and Glimcher, 2005; Daw *et al.*, 2006; Bayer *et al.*, 2007; Li *et al.*, 2006).

We argue that this approach, while providing compelling evidence that dopamine is worthy of further study, is not the best way of testing the dopaminergic hypothesis, for four related reasons. First, it is clear that any test of the RPE model derived in this way must be a joint test of both the RPE hypothesis *and* the proposed relationship between reward, beliefs, and the observable world. For example, the RPE model could be completely accurate, but the way in which beliefs are formed could be very different from that in the proposed model under test. Under these circumstances, the current tests could incorrectly reject the RPE hypothesis.

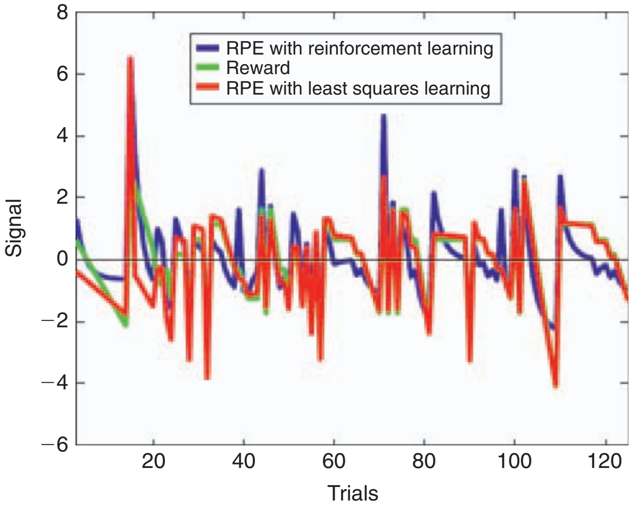


FIGURE 3.1 Estimated signals generated from simulations of the experiment in Li *et al.* (2006): Taking the experimental design reported in this paper, we simulate an experimental run, and calculate the output of various transforms of the resulting sequence of rewards. The graph shows the path of reward itself, a reward prediction error signal calculated from a reinforcement learning model and a reward prediction error signal calculated with a least-squares model of learning.

Second, such an approach can make it very difficult successfully to compare and contrast different models of dopamine activity, as the models themselves are poorly defined. If, for example, it were found that a certain data set provided more support for the RPE hypothesis than the salience hypothesis, a committed follower of the salience school could claim that the problem is in the definition of reward or salience. Given enough degrees of freedom, such a person could surely come up with a definition of salience which would fit the provided data well. Thus, tests between hypotheses can descend into tests of specific parametric specifications for "salience" or "reward."

Third, this can lead in practice to tests which do not have a great deal of power to differentiate between different hypotheses. Figure 3.1 shows the path of three different variables calibrated on the experimental design of Li *et al.* (2006): RPE as calculated by the authors, reward unadjusted by expectations, and RPE using a least-squares learning rule. It is obvious that these three lines are almost on top of each other. Thus, the fact that calculated RPE is correlated with brain activity is not evidence that such an area is encoding RPE; the RPE signal would also be highly correlated with any brain area which was encoding reward – or indeed one which just kept track of the amount of money available.

Fourth, the technique usually employed to solve such problems, which is to run statistical "horse races" between different models, is in itself problematic: statistical tests of non-nested models are themselves controversial. The "degrees of freedom" problem discussed above makes it very difficult to discount a particular model, as the model may be adapted so as better to fit the specific data. And even if it is shown that a particular model fits better than another, all this tells us is that the model we have is the best fitting of those considered. It doesn't tell us that the model is better than another model that we haven't thought of, or that the data don't deviate from our proposed model in some important, systematic way.

Because of these problems, we take an alternative, axiomatic approach to modeling RPE. Just as with utility theory, this approach is completely agnostic regarding how latent variables are related to other variables in the outside world. Instead, these variables are identified only in relation to their effect on the object of interest – in this case, dopamine. We ask the following question: Say that there is such a thing as "reward" which people assign to different objects or experiences, and "beliefs" (or expectations) which they assign to different circumstances, and dopamine responds to the difference between the two: what are the properties that dopamine activity must obey? In other words, when can we find *some* definition of rewards and some definition of expectation such that dopamine responds to the difference between the two? The resulting theory takes the form of a set of behavioral rules, or axioms, such that the data obey the RPE model if, and only if, these rules are satisfied. The problem of jointly testing the RPE theory and the definition of reward and belief is solved by defining both concepts *within* the theory, and only in relation to dopamine.

Our axioms enable us to characterize the entire class of RPE models in a simple, non-parametric way, therefore boiling the *entire* class of *RPE models* down to its essential characteristics. The axioms tell us exactly what such models imply for a particular data set – nothing more and nothing less. Hence our tests are *weaker* than those proposed in the traditional method of testing the RPE hypothesis described above. We ask only whether there is some way of defining reward and expectations so as to make the RPE model work. The traditional model in addition demands that rewards and beliefs are of a certain parametric form. Our tests form a basic minimal requirement for the RPE model. If the data fail our tests, then there is no way that the RPE model can be right. Put another way, if brain activity is to satisfy any one of the entire class of models that can be tested with the "traditional" approach, it must also satisfy our axioms. If dopaminergic responses are too complicated to be explained by our axioms, then, *a fortiori*, they are too complex to be fit using standard models of reward prediction error learning. Moreover, our approach allows us to perform hierarchical tests of a particular model – starting with the weakest possible formulation, then testing increasingly structured variants to find out what the data will support. A final and related point is that it allows for constructive interpretation of failures of the model. By knowing which axiom is violated, we can determine how the model-class must be adjusted to fit the data.

In order to provide the cleanest possible characterization, we develop the RPE model in the simplest environment in which the concept of a reward prediction error makes sense. The agent is endowed a lottery from

BOX 3.1

A GLOSSARY OF TERMS

Here, we provide a guide to the terms and symbols from economics used in describing the RPE model and its axiomatic basis:

Prize: One of the objects that a subject could potentially receive (e.g. amounts of money, squirts of juice) when uncertainty is resolved.

Lottery: A probability distribution over prizes (e.g., 50% chance of winning \$5, 50% chance of losing \$3).

Support: The set of prizes that can potentially be received from a lottery (e.g., for the lottery 50% chance of winning \$5, 50% chance of losing \$3, the support is {\$5, −\$3}).

Degenerate lottery: A lottery with a 100% probability of winning one prize.

$\in$: "is a member of" in set notation (e.g., $x \in X$ indicates that x is an element of the set X, or "New York" $\in$ "American cities").

$\mathbb{R}$: The set of all real numbers.

$\rightarrow$: "mapping to," used to describe a function, so $f : X \rightarrow Y$ indicates a function f which associates with each element in set X a unique element in set Y.

$|$: "objects that satisfy some condition" – for example, $\{(z, p) \mid z \in Z, p \in \Lambda(z)\}$ means any z and p such that z is an element of Z and p is an element of $\Lambda(z)$.

which a prize is realized. We observe the dopaminergic response when each possible prize z is realized from lottery p, as measured by the *dopamine release function*. Many of the mathematical subtleties of the theory that follow derive from the fact that it is not possible to observe dopaminergic responses to prizes that are not in the support of a particular lottery[6].

Definition 1

The set of prizes is a metric space Z with generic element $z \in Z$[7]*. The set of all simple lotteries (lotteries with finite support) over Z is denoted* Λ*, with generic element* $p \in \Lambda$*. We define* $e_z \in \Lambda$ *as the degenerate lottery that assigns probability 1 to prize* $z \in Z$ *and the set* $\Lambda(z)$ *as all lotteries with z in their support,*

$$\Lambda(z) \equiv \{p \in \Lambda | \, p_z > 0\}.$$

The function $\delta(z, p)$ *defined on* $M = \{(z, p) \,|\, z \in Z, p \in \Lambda(z)\}$ *identifies the dopamine release function,* $\delta : M \rightarrow \mathbb{R}$*.*

The RPE hypothesis hinges on the existence of some definition of "predicted reward" for lotteries and "experienced reward" for prizes which captures all the necessary information to determine dopamine output. In this case, we make the basic rationality assumption that the expected reward of a degenerate lottery is equal to its experienced reward as a prize. Hence the function $r : \Lambda \rightarrow \mathbb{R}$ which defines the expected reward associated with each lottery simultaneously induces the reward function on prizes $z \in Z$ as $r(e_z)$. We define $r(Z)$ as the set of values taken by the function r across degenerate lotteries,

$$r(Z) = \{r(p) \in \mathbb{R} | \, p = e_z, z \in Z\}.$$

What follows, then, are our three basic requirements for the DRPE hypothesis. Our first requirement is that there exists some reward function containing all information relevant to dopamine release. We say that the reward function fully summarizes the DRF if this is the case. Our second requirement is that the dopaminergic response should be strictly *higher* for a more rewarding prize than for a less rewarding one. Furthermore, a given prize should lead to a *higher* dopamine response when obtained from a lottery with *lower* predicted reward. Our third and final requirement is that, if expectations are met, the dopaminergic response does not depend on what was expected. If someone knows for sure that he is going to receive a particular prize, then dopamine must record that there is no "reward prediction error," regardless of how good or bad is the prize might be. We refer to this property as "no surprise constancy." These requirements are formalized in the following definition.

Definition 2

A dopamine release function $\delta : M \rightarrow \mathbb{R}$ *admits a dopaminergic reward prediction error (DRPE) representation if there exist a reward function* $r : \Lambda \rightarrow \mathbb{R}$ *and a function* $E : r(Z) \times r(\Lambda) \rightarrow \mathbb{R}$ *that:*

1. Represent *the DRF: given (z,p)* $\in M$*,*

$$\delta(z, p) = E(r(e_z), r(p)).$$

2. Respect dopaminergic dominance: *E is strictly increasing in its first argument and strictly decreasing in its second argument.*
3. *Satisfy* no surprise constancy: *given* $x, y \in r(Z)$*,*

$$E(x, x) = E(y, y).$$

We consider this to be the weakest possible form of the RPE hypothesis, in the sense that anyone who believes dopamine encodes an RPE would agree that it must have *at least* these properties. In Caplin and Dean (2008b) we consider various refinements, such as the case in which dopamine literally responds to the algebraic difference between experienced and predicted reward (i.e $\delta(z, p) = F(r(e_z) - r(p))$) and the case in which predicted reward is the mathematical expectation of experienced rewards (i.e $r(p) = \sum_{z \in Supp(p)} p(z) r(e_z)$). Both of these represent much more specific refinements of the DRPE hypothesis.

It turns out that the main properties of the above model can be captured in three critical axioms for $\delta : M \rightarrow \mathbb{R}$. We illustrate these axioms in Figures 3.2–3.4 for the two-prize case in which the space of lotteries Λ can be represented by a single number: the probability of winning prize 1 (the probability of winning prize 2 must be 1 minus the probability of winning prize 1). This forms the x-axis of these figures. We represent the function δ (i.e. dopamine activity) using two lines – the dashed line indicates the amount of dopamine released when prize 1 is obtained from each of these lotteries (i.e. $\delta(z_1, p)$), while the solid line represents the amount of dopamine released when prize 2 is obtained from each lottery (i.e. $\delta(z_2, p)$). Note that there are no observations at $\delta(z_1, 0)$ and $\delta(z_2, 1)$, as prize 1 is not in the support of the former, while prize 2 is not in the support of the latter.

---

[6] Caplin and Dean (2008b) covers the case in which lotteries are initially chosen from a set, and relates the reward representation below to the act of choosing.

[7] A metric is a measure of the distance between the objects in the space.

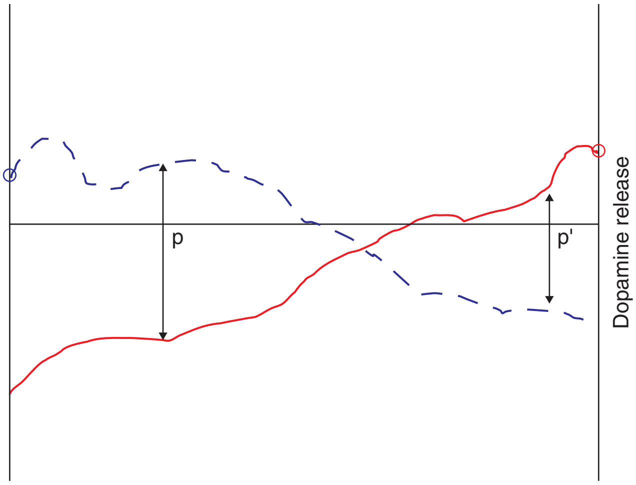


FIGURE 3.2 A violation of A1: when received from lottery *p*, prize 1 leads to higher dopamine release than does prize 2 indicating that prize 1 has higher experienced reward. This order is reversed when the prizes are realized from lottery *p′*, suggesting prize 2 has higher experienced reward. Thus a DRPE representation is impossible whenever the two lines cross.

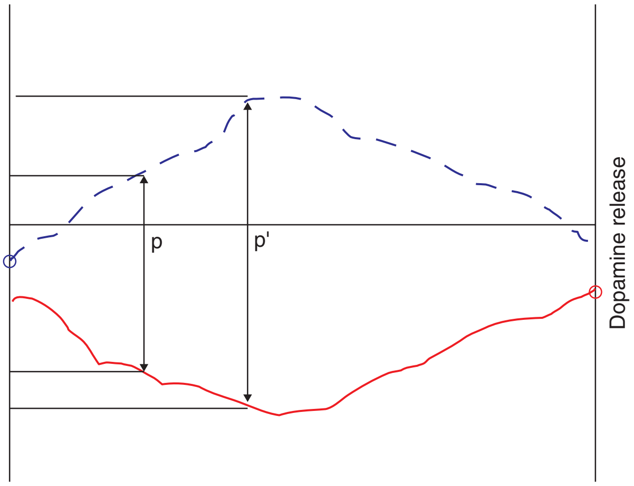


FIGURE 3.3 A violation of A2: Looking at prize 1, more dopamine is released when this prize is obtained from *p′* than when obtained from *p*, suggesting that *p* has a higher predicted reward than *p′*. The reverse is true for prize 2, making a DRPE representation impossible. This is true whenever the two lines have a different direction of slope between two points.

Our first axiom demands that the order on the prize space induced by the DRF is independent of the lottery that the prizes are obtained from. In terms of the graph in Figure 3.2, if dopaminergic release based on lottery *p* suggests that prize 1 has a higher experienced reward than prize 2, there should be no lottery *p′* to which dopaminergic release suggest that prize 2 has a higher experienced reward that prize 1. Figure 3.2 shows a violation of such *Coherent Prize Dominance.* It is intuitive that all such violations must be ruled out for a DRPE to be admitted. Our second axiom ensures that the ordering of lotteries by dopamine release is independent of the obtained prize. Figure 3.3 shows a case that contradicts this, in which more dopamine is released when prize 1 is obtained from lottery *p* than when it is obtained from lottery *p′*, yet the exact opposite is true for prize 2. Such an observation clearly violates the DRPE hypothesis. Our final axiom deals directly with equivalence among situations in which there is no surprise, a violation of which is recorded in Figure 3.4, in which more dopamine is released when prize 2 is obtained from its degenerate lottery (i.e. the lottery which gives prize 2 for sure) than when prize 1 is obtained from its degenerate lottery.

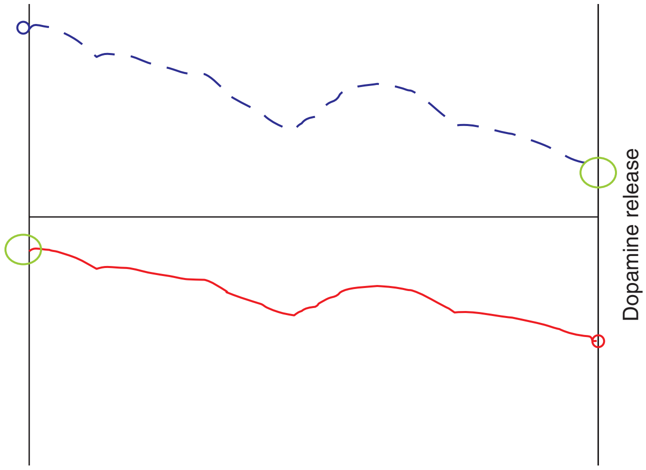


FIGURE 3.4 A violation of A3: the dopamine released when prize 1 is obtained from its sure thing lottery is higher that that when prize 2 is obtained from its sure thing lottery.

Formally, these axioms can be described as follows:

Axiom 1 (A1: Coherent Prize Dominance)
Given $(z, p),(z', p'),(z', p),(z, p') \in M$,

$$\delta(z, p) > \delta(z', p) \Rightarrow \delta(z, p') > \delta(z', p')$$

Axiom 2 (A2: Coherent Lottery Dominance)
Given $(z, p),(z', p'),(z', p),(z, p') \in M$,

$$\delta(z, p) > \delta(z, p') \Rightarrow \delta(z', p) > \delta(z', p')$$

Axiom 3 (A3: No Surprise Equivalence)
Given $z, z' \in Z$,

$$\delta(z', e_{z'}) = \delta(z, e_z)$$

These axioms are clearly necessary for any RPE representation. In general, they are not sufficient (see Caplin *et al.* (2008a) for a discussion of why, and what additional axioms are required to ensure an RPE representation). However, it turns out that these three axioms *are* sufficient in the case in which there are only two prizes (i.e. $|Z| = 2$). For a more general treatment of the problem, see Caplin and Dean (2008b) and Caplin *et al.* (2008a).

Notice how these axioms allow us to perform a clean, non-parametric test of the RPE hypothesis, without having to specify some auxiliary models for how rewards are related to prizes, and how beliefs (or reward expectations) are formed. The only assumption we make is that the "rewarding nature" of prizes, and the beliefs attached to each lottery, are consistent.

Our tests allow us to differentiate the RPE model from other models of dopamine activity: while A1–A3 form crucial underpinnings for the RPE hypothesis, they appear inconsistent with alternative hypotheses relating dopamine to salience (e.g. Zink *et al.*, (2003), and to experienced reward (e.g. Olds and Milner, (1954). Consider two prizes z and z', and two lotteries, p, which gives a 1% chance of winning z and a 99% chance of winning z', and p' which reverses these two probabilities. It is intuitive that that receiving z from p would be a very "salient," (or surprising) event, where as receiving z' would be very unsurprising. Thus a system responding to salience should give higher readings when z is obtained from p than when z' is obtained from p. However, this situation is reversed when the two prizes are obtained from p'. Thus we would expect A1 to fail if dopamine responded to salience. A similar argument shows that A2 would also fail, while A3 would hold, as the salience of getting a prize from a sure-thing lottery should be the same in all cases. With regard to the older theory that dopamine responds only to "experienced reward," this would lead A3 to be violated – different prizes with different reward values would give rise to different dopaminergic responses, even when received from degenerate lotteries.

In Caplin *et al.* (2008a) we describe the methodology by which we test the axioms described above. Essentially, we endow subjects with lotteries with varying probabilities (0, 0.25, 0.5, 0.75, 1) of winning one of two prizes (−\$5, \$5). We then observe brain activity using an fMRI scanner when they are informed of what prize they have won for their lottery. We focus on the nucleus accumbens, an area of the brain which are rich in dopamine output. While observing activity in this area is clearly not the same as observing dopamine, other authors (e.g., O'Doherty *et al.*, 2003, 2004; Daw *et al.*, 2006) claim to have found RPE-like signals using a similar technique. The noisy nature of fMRI data does, however, force us to confront the issue of how the continuous and stochastic data available to neuroscientists can be used to test axiomatic models. This is an area greatly in need of systemization. Caplin *et al.* (2008a) take the obvious first step by treating each observation of fMRI activity when some prize p is obtained from some lottery z as a noisy observation of actual dopamine activity from that event. By repeated sampling of each possible event, we can used standard statistical methods to test whether we can reject the null hypothesis that, for example, $\delta(p, z) = \delta(q, w)$ against the hypothesis that $\delta(p, z) > \delta(q, w)$. It is these statistical tests to test the axioms that form the basis of our theory.

CONCLUSIONS

The results reported in *Caplin et al.* (2008a) suggest that we can indeed identify areas of the brain whose activity is in line with the basic RPE model. We can therefore begin to refine our model of dopamine activity, for example by deepening our understanding of how reward assessments vary with beliefs. In Caplin and Dean (2008b), we illustrate this process with an extreme example in which beliefs must be equal to the mathematical expectation of experienced rewards. A further step is to introduce models of subjective beliefs and learning to the RPE model, a direction of expansion required to capture the hypothesized role of dopamine in the process of reinforcement learning. Once we have completed these initial experiments, we intend to use the apparatus to start addressing questions of economic importance – exploring the use of dopaminergic measurements to open a new window into the beliefs of players in game theoretic settings and to understand addictive behavior (an endeavor already begun by Bernheim and Rangel, 2004).

In practical terms, improvements in measurement technology will be vital as we refine our axiomatic model. For that reason we are intrigued by the measurement techniques pioneered by Phillips and colleagues (2003), and others, that are enabling dopaminergic responses to be studied ever more closely in animals. The increased resolution that these techniques makes possible may enable us to shed an axiomatic light on whether or not dopamine neurons are asymmetric in their treatment of positive than negative reward prediction errors, as conjectured by Bayer and Glimcher [2005]. Axiomatically inspired experimentation may also allow progress to be made on whether or not signals of reward surprise may be associated with neurons that are associated with different neurotransmitters, such as serotonin.

Our axiomatic approach to neuroeconomics forms part of a wider agenda for the incorporation of nonstandard data into economics. Recent advances in experimental techniques have led to an explosion in the range of data available to those interested in decision making. This has caused something of a backlash within economics against the use of non-standard data in general and neuroscientific data in particular. In their impassioned defense of "mindless economics," Gul and Pesendorfer (2008) claim that non-choice data cannot be used as evidence for or against economic models, as these models are not designed to explain such observations. By design, our axiomatic approach is immune to such criticisms as it produces models which formally characterize whatever data is under consideration. In a separate sequence of papers, we apply the same approach to a data set which contains information on how choices change over time (Caplin and Dean, 2008a; Caplin *et al.*, 2008b). We show how this expanded data set can give insight into the process of information search and choice.

Ideally, an expanded conception of the reach of the axiomatic methodology will not only open new directions for neuroeconomic research, but also connect the discipline more firmly with other advances in the understanding of the process of choice, and the behaviors that result.

References

Allais, M. (1953). Le comportement de l'homme rationnel devent de risque: critique des postulates et axiomes de l'ecole Americaine. *Econometrica* 21, 503–546.

Bayer, H. and Glimcher, P. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141.

Bayer, H., Lau, B., and Glimcher, P. (2007). Statistics of midbrain dopamine neuron spike trains in the awake primate. *J. Neurophysiol.* 98, 1428–1439.

Berridge, K.C. and Robinson, T.E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Rev.* 28, 309–369.

Bernheim, B.D. and Rangel, A. (2004). Addiction and cue-triggered decision processes. *Am. Econ. Rev.* 94, 1558–1590.

Caplin, A. and Dean, M. (2008a). *The choice process*. Mimeo, New York University.

Caplin, A. and Dean, M. (2008b). Dopamine, reward prediction error, and economics. *Q. J. Economics* 123(2), 663–701.

Caplin, A., Dean, M., Glimcher, P., and Rutledge, R. (2008a). Measuring beliefs and Rewards: A neuroeconomic Approach. Working Paper, New York University.

Caplin, A., Dean, M., and Martin, D. (2008b). *The choice process: experimental evidence*. Mimeo, New York University.

Daw, N., O'Doherty, J.P., Dayan, P. *et al*. (2006). Polar exploration: cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.

Ellsberg, D. (1961). Risk, ambiguity and the savage axioms. *Q. J. Economics* 75, 643–669.

Gardner, E. and David, J. (1999). The neurobiology of chemical addiction. In: J. Elster and O.-J. Skog (eds), *Getting Hooked: Rationality and Addiction*. Cambridge: Cambridge University Press, pp. 93–115.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with a non-unique prior. *J. Math. Econ.* 18, 14–153.

Gul, F. (1991). A theory of disappointment aversion. *Econometrica* 59, 667–686.

Gul, F. and Pesendorfer, W. (2008). The case for mindless economics. In: A. Caplin and A. Schotter (eds), *Handbook of Economic Methodology, Vol. 1, Perspectives on the Future of Economics: Positive and Normative Foundations*. Oxford: Oxford University Press, pp. 3–40.

Kahneman, D. and Tversky, A. (1973). On the psychology of prediction. *Psychol. Rev.* 80, 237–25l.

Kreps, D.M. and Porteus, E.L. (1979). Dynamic choice theory and dynamic programming. *Econometrica* 47(1), 91–100.

Kiyatkin, E.A. and Gratton, A. (1994). Electrochemical monitoring of extracellular dopamine in nucleus accumbens of rats lever-pressing for food. *Brain Res.* 652, 225–234.

Li, J., McClure, S.M., King-Casas, B., and Montague, P.R. (2006). Policy adjustment in a dynamic economic game. *PlosONE* 1(1), e103.

Lucas, R. (1971). Asset prices in exchange economy. *Econometrica* 46(6), 1429–1445.

Mirenowicz, J. and Schultz, W. (1994). Importance of unpredictability for reward responses in primate dopamine neurons. *J. Neurophysiol.* 72(2), 1024–1027.

Montague, P.R. and Berns, G.S. (2002). Neural economics and the biological substrates of valuation. *Neuron* 36, 265–284.

Montague, P.R, Dayan, P. and Sejnowski, T.J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. NeuroSci*. 16: 1936–1947.

Olds, J. and Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *J. Comp. Physiol. Psychol.* 47, 419–427.

O'Doherty, J., Dayan, P., Friston, K.J. *et al*. (2003). Temporal difference models account and reward-related learning in the human brain. *Neuron* 38, 329–337.

O'Doherty, J., Dayan, P., Schultz, J. *et al*. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454.

Phillips, P.E., Stuber, G.D., Heien, M.L. *et al*. (2003). Subsecond dopamine release promotes cocaine seeking. *Nature* 422(6932), 614–618.

Pratt, J. (1964). Risk aversion in the small and the large. *Econometrica* 32, 122–136.

Quiggin, J. (1982). A theory of anticipated utility. *J. Econ. Behav. Org.* 3, 323–343.

Redgrave, P. and Gurney, K.N. (2006). The short-latency dopamine signal: a role in discovering novel actions? *Nat. Rev. Neurosci.* 7, 967–975.

Samuelson, P. (1938). A note on the pure theory of consumer's behavior. *Economica* 5, 61–71.

Schmeidler, D. (1982). Subjective probability without additivity. Working Paper, The Foerder Institute for Economic Research, Tel Aviv University.

Schultz, W., Apicella, P., and Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.* 13, 900–913.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*, 1953 edn. Princeton, NJ: Princeton University Press.

Zink, C.F., Pagnoni, G., Martin, M.E. *et al*. (2003). Human striatal response to salient nonrewarding stimuli. *J. Neurosci.* 23, 8092–8097.

CHAPTER

4

Neuroeconomics: Formal Models of Decision Making and Cognitive Neuroscience

Aldo Rustichini

OUTLINE

INTRODUCTION

We provide here a link between the formal theory of decision making and the analysis of the decision process as developed in neuroscience, with the final purpose of showing how this joint analysis can provide explanation of important elements of economic behavior.

The methodological standpoint we present is that experimental economics, including neuroeconomics, establishes relationships among variables, some inferred from observed behavior. In particular, a fundamental component of the neuroeconomics project is to establish connections between variables derived from observed behavior and psycho-physiological quantities. For example, the derived variables can be utility, or parameters like risk aversion. If a researcher claims that the relationship between utility or value and a psycho-physiological quantity (like firing rate of a neuron) is linear, then one has to be sure that the derived variable is uniquely defined, up to linear transformations. If it is not, the statement is meaningless.

We first review the basic concepts and results in decision theory, focusing in particular on the issue of cardinal and ordinal utility, remembering that within the von Neumann-Morgenstern framework the utility function on lotteries (the only observable object) is only defined up to monotonic transformations.

We then show how, under well-specified assumptions, it is possible to identify a unique ordinal object, and how this is based on stochastic choice models. These are, however-static models, so they do not give an account of how the choice is reached. We show that the static models have a dynamic formulation, which extends the static one.

Once the main features of the decision process have been established, we can show how they explain important features of the choice – even some that had been ignored so far. For example, we show how risk aversion, impatience, and cognitive abilities are related.

AXIOMATIC DECISION THEORY

In economic analysis, decision theory is developed with a purely axiomatic method. The theory proceeds by first defining a set of choices that a subject (the decision maker, DM) faces. A choice is a finite set of options that are offered to the DM; a decision is the selection of one of these options. The observed data are pairs of choices offered and decisions taken: it is possible to collect these data experimentally asking a real DM to pick one out of two options, under the condition that the object selected is actually delivered to her.

The method and the main results of the theory are best illustrated in a simple and concrete example of choice environment, choice under risk. In this environment, the options are lotteries. A common lottery ticket provides an example of the abstract concept of lottery: a winning number is drawn at random, and with such a ticket, a person is entitled to a payment if the winning number is the one she has, and she receives no payment otherwise. In general, a lottery is a contract specifying a set of outcomes (the payments made to the subject in our example) and a probability for each of these outcomes. The probability is specified in advance and known to the subject, so in this model there is only objective uncertainty, as opposed to the subjective uncertainty analyzed in Savage (1954) and Anscombe and Aumann (1963).

A lottery with two outcomes can be formally described with a vector $(x, p, y, 1-p)$, to be interpreted as: this lottery gives the outcome x with probability p, and the outcome y with probability $1-p$. For example, the lottery ($10, 1/2, $0, 1/2) where outcomes are monetary payments gives a 50-50 chance of a payment of $10, and nothing otherwise. Lotteries do not need to be a monetary amount, but for simplicity of exposition we confine ourselves to this case.

The Method of Revealed Preferences

We can observe the decisions made by our subject, while we do not observe her preferences directly. However, we may interpret her choices as a "revelation" that she makes of her preferences. Suppose that when she is presented with a choice between lottery L_1 and L_2 she chooses L_1: we may say that she reveals she prefers L_1 to L_2. Within economic analysis, it is in this sense, and in this sense only, that we can say that the DM prefers something. The two descriptions of her behavior, one with the language of decisions and the other with that of preferences, are, by the definition we adopt, perfectly equivalent. Since the language of preferences seems more intuitive, it is the one used typically by decision theory, and is the one used here. But how do we describe the behavior, or preferences, of our subject?

Axioms

Even with simple lotteries with two monetary outcomes, by varying the amounts and the probabilities we can obtain an infinite set of possible lotteries, and by taking all the possible pairs of these lotteries we can obtain infinitely many choices. To describe the behavior of a subject completely, we should in principle list the infinite set of decisions she makes. To be manageable, a theory needs to consider instead subjects whose decisions can be described by a short list of simple principles, or axioms.

The first axiom requires that the preferences are complete: for every choice between the two lotteries L_1 and L_2, either L_1 is preferred to L_2, or L_2 is preferred to L_1. The occurrence of both possibilities is not excluded: in this case, the subject is indifferent between the two lotteries. When the subject prefers L_1 to L_2, but does not prefer L_2 to L_1, then we say that she strictly prefers L_1 to L_2. The second axiom requires the preferences to be transitive: if the DM prefers L_1 to L_2 and L_2 to L_3, then she prefers L_1 to L_3. We define the preference order $\succeq$ by writing $L_1 \succeq L_2$ when decision maker prefers L_1 to L_2 and we write $L_1 \succ L_2$ when decision maker strictly prefers L_1 to L_2. Formally:

Axiom 1 (Completeness and transitivity)

For all lotteries L_1, L_2 and L_3

1. *Either $L_1 \succeq L_2$ or $L_2 \succeq L_1$*
2. *If $L_1 \succeq L_2$ and $L_2 \succeq L_3$ then $L_1 \succeq L_3$.*

The next two axioms are also simple, but more of a technical nature. Suppose we have two lotteries, $L_1 = (x, p, y, 1 - p)$ and $L_2 = (z, q, w, 1 - q)$. Take any number r between 0 and 1. Imagine the following contract. We will run a random device, with two outcomes, *Black* and *White*, the first with probability r. If *Black* is drawn, then you will get the outcome of the lottery L_1; if *White* is drawn you will get the outcome of the lottery L_2. This new contract is a compound lottery. If you do not care about how you get the amounts of money, then this is the lottery with four outcomes described as $(x, rp, y, r(1 - p), z, (1 - r)q, w, (1 - r)(1 - q))$. We write this new lottery as $rL_1 + (1 - r)L_2$.

The next axiom requires that if you strictly prefer L_1 to L_2, then for some number r, you strictly prefer $rL_1 + (1 - r)L_2$ to L_2. This seems reasonable: when r is close to 1 the composite lottery is very close to L_1, so you should strictly prefer it to L_2 just like you strictly prefer L_1.

Axiom 2 (Archimedean continuity)

If $L_1 \succ L_2$ then for some number $r \in (0,1)$,

$$rL_1 + (1 - r)L_2 \succ L_2.$$

Finally, suppose that you strictly prefer L_1 to L_2. Then for any lottery L_3, you also strictly prefer $rL_1 + (1 - r)L_3$ to $rL_2 + (1 - r)L_3$. Again, this seems reasonable. When, in the description we gave above, *White* is drawn, then in both cases you get L_3; when *Black* is drawn, in the first case you get L_1 and in the second L_2. Overall, you should prefer the first lottery $rL_1 + (1 - r)L_3$.

Axiom 3 (Independence)

If $L_1 \succ L_2$ then for any number $r \in (0,1)$ and any lottery L_3,

$$rL_1 + (1 - r)L_3 \succ rL_2 + (1 - r)L_3.$$

Representation of Preferences

A fundamental result in decision theory (due to von Neumann and Morgenstern, (vNM), 1947) is that subjects having preferences that satisfy these axioms (completeness, transitivity, Archimedean continuity and independence) behave as if they had a simple numerical representation of their preferences – that is, a function that associates with a lottery a single number, called the utility of the lottery, that we can write as $U(L)$. This function is called a representation of the preferences if whenever L_1 is preferred to L_2, then the utility of L_1 is larger than the utility of L_2, that is $U(L_1) > U(L_2)$. (Note that here we use $>$ not $\succ$ because $U(L_1)$ is a numerical property not a preference.)

The vNM theorem also states that the preference order satisfies the axioms above if, and only if, the numerical representation has a very simple form, equal to the expectation of the utility of each outcome, according to some function u of outcomes. For example, the expected utility of the lottery $L = (x, p, y, 1 - p)$ is:

$$U(L) = pu(x) + (1 - p)u(y) \tag{4.1}$$

Cardinal and Ordinal Utilities

For neuroeconomics, and any research program that tries to determine how decisions are implemented, the utility function is the most interesting object. This function ties observed behavior with a simple one-dimensional quantity, the utility of the option, and predicts that the decision between two options is taken by selecting the option with the highest utility. However, if we are interested in determining the neural correspondents of the objects we have introduced, we must first know whether these objects are unique. For example, we may formulate the hypothesis that the decision is taken depending on some statistics of the firing rate of a group of neurons associated with each of the options. We may also consider that this firing rate is proportional to the utility we determine from observed choice behavior. Then we need to know whether this utility is uniquely determined. This introduces us to a fundamental distinction in decision theory, between cardinal and ordinal representation.

An ordinal representation of a preference is any utility function such that $U(L_1) > U(L_2)$ if, and only if, L_1 is strictly preferred to L_2. There are clearly many such functions. For example, if M is any strictly increasing function, then also $M(U(L_1)) > M(U(L_2))$ if, and only if, L_1 is strictly preferred to L_2. So we say that an ordinal representation is only unique up to increasing (or monotonic) transformation, like the one we have used from U to $M(U)$. Consider now the function u is equation (4.1), and take two numbers $a > 0$ and b. Replace the u function in (4.1) with the new function v defined for any value z by $v(z) = au(z) + b$. If we replace the u in equation (4.1) we obtain a new function on lotteries which also represents preferences, and has the form of expected utility. Since these transformations leave the observed choices and preferences unchanged, the u in (4.1) is not unique.

However, these are the only transformations we can apply. A second remarkable part of the vNM theorem is that if two functions u and v represent the preferences of a subject as expected utility (that is, as in (4.1)), then it must be that $v(z) = au(z) + b$ for some positive number a and some number b. In this case the two functions are said to be linear transformations of each other, and representations like these are called cardinal representations. A different but equivalent way of saying this is that if we consider functions on a range of monetary prizes between a minimum of 0, say, and a maximum value M, and we agree to normalize the utility function u to $u(0) = 0$ and $u(M) = 1$, then there is a unique such function that, once substituted in equation (4.1), represents the preferences of the DM.

However, the observed decision between two choices is determined by the function U, and this is only unique up to *monotonic* (not necessarily linear) transformations. So even if we agree to normalize $U(0) = 0$ and $U(M) = 1$, there are still infinitely many such Us. If we are looking for a neural basis of choice, then the only sensible statements that involve the function U are those that remain true if we take monotonic transformations of that function. For example, statements like "the firing rate is a linear transform of the U" are meaningless.

Can we do better than this? We can, if we agree to extend the set of observed data to include errors and time in the decision process. This will take us to the next topic of stochastic choices, and one step closer to the models of decision currently applied in neuroscience.

STATIC STOCHASTIC CHOICE

To illustrate and motivate this new point of view, we begin with a finding discovered in the 1940s by an Iowa researcher, D. Cartwright (Cartwright, 1941a, 1941b; Cartwright and Festinger, 1943). He asked subjects to pick one of two alternatives. By changing the parameter appropriately, the experimenter could make the choice more or less difficult – for example, setting the width of two angles closer would make the task of choosing the wider angle between the two a more difficult task. Also, by asking the subject to make the same choice repeatedly, at some distance in time, he could test the frequency of the choice of one or the other of the alternatives in different decision problems. He could now construct what we can call the empirical random choice: for every set of options, the frequency of choice of each option out of that set.

He also measured the response time for each choice and then plotted the average response time for each decision problem against the minimum frequency of any of the two choices in that same problem. The key finding was that the longest response time was observed when the minimum frequency was approaching 50%; the problems in which the subject was more likely to select, in different trials, both options were also those in which she was taking more time to decide. A related result is the "symbolic distance" effect, first stated in Moyer and Landauer (1967).

The finding of Cartwright suggests a model of decision where two opposing forces push in the direction of each of the options. When the difference between these two forces is large, the decision is frequently in favor of the favored option, and the decision is taken quickly. When they are the same, the frequency of choice of the two options becomes closer, and the response time becomes longer.

For our purposes of outlining a theory of the decision process when the decision is among economic choices, it is important to note that for economic choices the same result holds. Suppose we determine the utility of a subject from the observed choices, that is, the quantity $U(L)$ for every lottery L. We can now measure the distance between the utility of any two lotteries in a choice, and conjecture that the analogue of the Cartwright results holds in this situation: the closer the two options in utility, the longer the time to decide, and the higher the minimum probability of choosing any of the two. This conjecture has been confirmed in several studies. There is one problem, however: what is the distance between the utilities? If the utility is unique up to monotonic affine transformations, then the distance is well defined up to re-scaling by a single number. But we have just seen that the U in (4.1) is not unique up to monotonic affine transformations, thus even after normalization we have infinitely many such functions. So how can we measure in a meaningful way the distance in utility between two options? The key to a solution is in the inconsistency of choice that we have just reported.

Economic Theories of Static Stochastic Choice

The experimental evidence reviewed in the previous section suggests that when repeatedly faced with a choice between the same two options, the subject may not always choose the same option in each instance. In contrast, the utility theory we have reviewed so far predicts that if the utility of one of the two is larger, that should always be the chosen one. The key idea of the stochastic theory of choice is that the relative frequency of the choice of one option over the other gives a better measure of the utility of the two options.

There are two classes of models of stochastic choice in economic theory. Both address the following problem. Suppose that a DM is offered, in every period, the choice of a set of lotteries, a menu. We observe her choices over many periods. For a given menu, the choices may be different in different periods, but we can associate for every menu the frequency of choices over that menu – that is, a probability distribution over the set. Both classes of models want to determine the underlying preference structure that produces this observed frequency.

Let us state formally the problem that we have just described. For every nonempty set Y, let $\mathcal{P}(Y)$ be the set of all finite subsets of Y, and $\Delta(Y)$ be the set of all probability measures over Y. Let X be a set of options: for example, the set of lotteries that we have considered so far. A *random choice rule* (*RCR*) σ is a function from $\mathcal{P}(X)$ to $\Delta(X)$, mapping an element $D \in \mathcal{P}(X)$ to σ^D, such that for every such D, $\sigma^D(D) = 1$. The value $\sigma^D(x)$ is the observed frequency of the choice of x out of D.

Random Utility Models

In random utility models (see McFadden and Richter, 1991, for an early axiomatic analysis, and Gul and Pesendorfer, 2003, for a very recent development) the subject has a set of different potential utility functions (almost different selves), and only one of them is drawn every time she has to make a decision. This momentarily dominant utility decides the choice for that period. Since utilities are different, the choices from the same set of options may be different in different times, although in every period the DM picks the best option.

The hypothesis that random choice is produced by random utilities imposes restrictions on observed behavior. For example, in this class of models choices are made from a set of lotteries, called a menu. Since each of these utility functions is linear, the choice is always in a special subset of the menu (technically, its boundary). A representation of the random-choice rule in random utility models is a probability distribution over utilities such that the frequency of the choice of x out of D, $\sigma^D(x)$ is equal to the probability of the set of utilities that have the element x as a best choice out of D.

Stochastic Choice Models

In stochastic choice models, the utility function is the same in every period. The DM does not always choose the option with the highest utility, but she is more likely to choose an option, the higher its utility is compared to that of the other options. The power of these models is based on two ideas. The first is the decomposition of the decision process in two steps; evaluation and choice. The second is that frequency of choice gives a measure of the strength of preferences. Together, they give a way to identify a cardinal utility. Early axiomatic analysis of this problem is in Davidson and Marschak (1959) and in Debreu (1958). A set of axioms that characterize *RCR*s which have a stochastic choice representation and that separate these two ideas is presented in, Maccheroni *et al.* (2007). We examine both ideas in detail.

Utility Function and Approximate Maximization

A representation in stochastic choice models has two elements. The first is the evaluation, which is performed by a utility function that associates a real number with each option in the available set. The second is an approximate maximization function associating to each vector of utilities the probability of choosing the corresponding option.

The utility function is naturally determined on the basis of the random choice rule σ. Write $\sigma(x, y) = \sigma^{\{x,y\}}(x)$ and consider the relation defined by

$$x \succeq y \text{ if and only if } \sigma(x, y) \geq \sigma(y, x).$$

As usual, a function u on X represents the order $\succeq$ if $x \succeq y$ if, and only if, $u(x) \geq u(y)$.

To define the second element, fix u and let U be the range of this function: $U = u(X)$. An approximate maximum selection is a function p from $\mathcal{P}(U)$ to $\Delta(U)$, associating with set A a probability p^A which is concentrated on A (that is, $p^A(A) = 1$) and is monotonic (that is, for every $a, b \in A$ if $a \geq b$ then $p^A(a) \geq p^A(b)$).

A representation of the *RCR* σ in stochastic choice models is given by a pair (u, p) of a utility function u on X representing $\succeq$ and an approximate maximization function p such that

$$\sigma^D(x) = p^{u(D)}(u(x)) \tag{4.2}$$

In Maccheroni *et al.* (2007) give a set of axioms that characterize *RCR* with such a representation.

Moreover, a pair (v, q) represents σ if, and only if, there exists an increasing function $g : u(X) \to R$ such that

$$v = g \circ u \quad \text{and} \quad q^B(b) = p^{g^{-1}(B)}(g^{-1}(b)) \quad \forall B \in \mathcal{P}(v(X)) \tag{4.3}$$

In other words, the function u is only determined up to monotonic (not just affine) transformations, so it is still an ordinal, not a cardinal object. Stochastic choice, by itself, does not imply the existence of and does not reveal a cardinal utility.

Strength of Preferences

A measure of the *strength of the preferences* of the DM indicates, for any x, y, z, and w, whether she prefers x to y more than she does z to w. As a special case, it also indicates whether she prefers x to y more than she does z to z itself – that is, whether she prefers x to y, so strength of preferences contains implicitly a preference order. How do we access this measure? One way is through verbal statements made by the DM: she introspectively evaluates the strength and communicates it to the experimenter, with words, not with choice.

Stochastic choice provides us with a second, objective way of measuring the strength of preferences. The value $\sigma(x, y)$ describes how frequently the option x is chosen instead of y. If we compare the frequency of choices out of two other options $\{z, w\}$, and we observe that $\sigma(x, y) > \sigma(z, w)$, then we may say that the DM likes x more than y with stronger intensity than she likes z more than she likes w: we write $(x, y) \geq (z, w)$ to indicate this order over pairs. A random choice rule as characterized in representation (4.2) is a measure of the strength preferences.

Representation (4.2) shows clearly that knowing the strength of preference does not by itself determine the utility function as a cardinal object. We can always introduce a monotonic transformation of the u function, provided we undo this transformation with an appropriate transformation of the approximate maximization function p.

To obtain u as a cardinal object, a specific and strong condition on the random choice rule is needed. The nature of the condition is clear: u is a cardinal object if the strength of preference only depends on the difference in utility, namely if the following difference representation holds:

$$\begin{gathered}\sigma(x,y) \geq \sigma(z,w) \text{ if, and only if,} \\ u(x) - u(y) \geq u(z) - u(w).\end{gathered} \tag{4.4}$$

Debreu (1958) investigates conditions insuring that condition (4.4) is satisfied. A necessary condition for the existence of a u as in (4.4) is clearly:

$$\begin{gathered}\sigma(x,y) \geq \sigma(x',y') \quad \text{and} \quad \sigma(y,z) \geq \sigma(y',z') \\ \text{imply } \sigma(x,z) \geq \sigma(x',z')\end{gathered} \tag{4.5}$$

(see also Krantz *et al.*, 1971; Shapley, 1975; Köbberling, 2006).

Together with an additional technical axiom (solvability), axiom (4.5) is all is needed for the existence of a function u that is a cardinal object: that is, if a function v also satisfies (4.4) then v is an affine monotonic transformation of u; that is, there are two numbers $a > 0$ and b such that $v = au + b$.

This opens the way for a complete stochastic choice representation of the random choice rule, with the additional condition that the utility u is cardinal. In a complete model of stochastic choice, if we introduce the additional axiom (4.5) then the approximate maximization function p depends only on the differences, that is:

$$p^{\{r,s\}}(r) = P(r - s) \tag{4.6}$$

for some function P. The question is now: how is that probability P implemented?

DYNAMIC STOCHASTIC CHOICE

In the plan of determining the neural basis of decision, we have two final steps. First, we have to produce a model of the decision process that produces a stochastic choice as described in the previous section. Second, we have to specify and test the neural basis implementing this process. Let us begin with the first.

The Random Walk Model

The model's original formulation is in Ratcliff (1978). As the title indicates, the theory was originally developed for memory retrieval, where the task is as follows. A subject has to decide whether an item that is in front of her is the same as one she has seen sometimes in the past, or not. She has the following information available. First, she has the visual evidence of the object in front of her. This object can be described abstractly as a vector of characteristics – the color, the smoothness of the surface, the width, the length, and so on. The subject also has some memory stored of the reference object, which can again be described by a vector of the same characteristics as the first one. If the description of the object is very detailed, the vector is a high-dimensional vector. The subject has to decide whether the object in front of her is the same as the object stored in memory, so she has a simple binary (*yes*, it is the same object, or *no*) decision to take.

In an experimental test, we can measure the time the subject takes to decide, her error rate, and how these variables depend on some parameter that we control – for example, how different the two objects are.

A plausible model of the process is as follows. The subject compares, one by one, each coordinate in the vector of characteristics of the real and recalled object. She may find that, to the best of her recollection, they coincide, or they do not. She proceeds to count the number of coincidences: an agreement of the features is taken as evidence in favor of "yes," a disagreement as evidence of "no". If the vector of evidence is very long, the subject may decide to stop before she has reviewed all the characteristics, according to a simple stopping rule: decide in favor of "yes" the first time the number of agreements minus the number of disagreements is larger than a fixed threshold; decide in favor of "no" when a similar lower barrier is reached. The general form of a decision process based on this idea is the random walk of decision. The model has been presented in a discrete or continuous time version. In the continuous time formulation, the process is typically assumed to be a Brownian motion, or at least a time homogeneous stochastic process.

The model has several parameters: first, those describing the process. For example, if the process is in continuous time and is a Brownian motion, the process is described by the mean and the variance. The second parameters are the barriers. There are at least two important observed variables: the probability that one of the two decisions is taken, and the time needed to reach the decision. The model has sharp predictions on the two variables: for example, if the drift in favor of one of the two options is stronger, then the probability of that option being chosen increases. Also, when the difference in drift between the two choice is small, then the time to take a decision increases.

DECISION IN PERCEPTUAL TASKS

Intense research regarding the neural foundation of the random walk model of decision has been undertaken in the past few years. To illustrate the method and the findings, we begin again with a classical experiment (Shadlen and Newsome, 1996, 2001; Schall, 2001).

In the experiment, the subject (for example, a rhesus monkey) observes a random movement of dots. A fraction of the dots is moving in one of two possible directions, left or right, while the others move randomly. The monkey has to decide whether the fraction of dots moving coherently is moving to the left or to the right, being instructed to do this after intensive training. If the monkey makes the right choice, it is compensated by a squirt of juice. Single neuron recording of neurons shows that the process of deciding the direction is the outcome of the following process: some neurons are associated with the movement to the left, and others to the right. The overall firing rate of the "left" and "right" neurons is, of course, roughly proportional to the number of dots moving in the two directions. The decision is taken when the difference between the cumulative firing in favor on one of the two alternatives is larger than a critical threshold.

Formal Model

A key feature of the information process described above is that each piece of information enters additively into the overall evaluation. This has the following justification. Suppose that information is about a state that is affecting rewards. A state is chosen by the experimenter, but is unknown to the subject. Information is provided, in every period, in the form of signals drawn independently in every period, from a distribution over signals that depends on the state. How is the information contained in the signal observed in every period aggregated?

In a simple formal example, suppose that the decision maker has to choose between two actions: left (l) and right (r). She receives a payment depending on the action she chooses and an unobserved state of nature $s \in \{L, R\}$; this is equal to $1 if, and only if, she chooses the correct action l if the state is L. Her utility is a function defined on the set $A \equiv \{l, r\}$ of actions and set of states $S \equiv \{L, R\}$ by $u(l, L) = u(r, R) = 1$, $(l, R) = u(r, L) = 0$. She has an initial subjective probability p that the state is R, and can observe a noisy signal on the true state of nature, according to the probability $P_s(x)$ of observing x at s.

The posterior odds ratio of L *versus* R with a prior P, after the sequence $(x_1, x_2, \ldots, x_n)$ is observed, is given by:

$$\frac{P(L|\, x_1, x_2, \ldots, x_n)}{P(L|\, x_1, x_2, \ldots, x_n)} = \frac{P(L)}{P(R)} \prod_{i=1}^{n} \left(\frac{P_L(x_i)}{P_R(x_i)} \right)$$

so that the log of the odds ratio are simply the sum of the log of the odd ratios of the signal

$$\log \frac{P(L|\, x_1, x_2, \ldots, x_n)}{P(L|\, x_1, x_2, \ldots, x_n)} = \log \frac{P(L)}{P(R)} + \sum_{i=1}^{n} \log \frac{P_L(x_i)}{P_R(x_i)}.$$

Decision in Economic Choices

We suggest that the mental operation that is performed when the subject has to choose between two economically valuable options consists of two steps. First, the individual has to associate a utility with each of the two options. Second, she then has to decide which of these two computed quantities is larger. This second step is a simple comparison of quantities. The first is completely new, and is specific to economic analysis. Note two important features of this model: first, even if the decision maker assigns (somewhere in her brain) a strictly larger utility to one of the two options, she still does not choose for sure that option: she only has a larger probability of doing so. Second, the decision maker has a *single* utility or preference order over outcomes. The choice outcome is not deterministic, because the process from utility evaluation to choice is random.

What is the evidence supporting this view? Let us begin from the step involving comparison of quantities. Experiments involving comparison of numbers, run with human subjects (see Sigman and Dehaene, 2005), confirm the basic finding that the response time is decreasing with the distance between the two quantities that are being compared. For example, if subjects have to decide whether a number is larger or smaller than a reference number, then the response time is decreasing approximately exponentially with the distance between the two numbers. So there is experimental evidence that suggests that the operation of comparing quantities follows a process that is close to that described by the random walk model. The last missing element is: do we have evidence that there are areas of the brain where neurons fire in proportion to the utility of the two options?

THE COMPUTATION OF UTILITY

In this experiment, a monkey is offered the choice between two quantities of different food or juices: for example, 3 units of apple, or 1 unit of raisin.

By varying the quantities of juice of each type offered, the experimenter can reconstruct, from "revealed preferences," the utility function of the monkey. This function can be taken to be, for the time being, an artificial construct of the theorist observing the behavior. The choices made by the subjects have the typical property of random choice: for example, between any amount less or equal to 2 units of apple and 1 unit of raisin, the monkey always chose the raisin. With 3 units of apple and 1 of raisin, the frequency of choice was 50/50 between the two. With 4 or more units of apple, the monkey always went for the apple. This is the revealed-preference evidence.

At the same time, experimenters can collect single neuron recording from areas that are known to be active in evaluation of rewards (for example, area 13 of the orbito-frontal cortex). They can then plot the average firing rate over several trials (on the y-axis) against the estimated utility of the option that was eventually chosen on the x-axis, thus obtaining a clear, monotonic relationship between the two quantities. These results are presented in detail in Chapter 29.

A Synthesis

We have now the necessary elements for an attempt to provide a synthesis of the two approaches, one based on economic theory and the other on neuroscience.

Consider a subject who has to choose between two lotteries. When considering each of them, she can assign to it an estimate of the expected utility of each option. This estimate is likely to be noisy. When she has to choose between the two lotteries, she can simply compare the (possibly noisy) estimate of the two utilities: thus the choice between the two lotteries is now determined by the comparison of these two values. At this stage, the choice is reduced to the task of comparing two numerical values, just as the task that the random walk model analyzes.

In summary, this model views the decision process as the result of two components: the first reduces the complex information describing two economic options to a numerical value, the utility of each option. The second performs the comparison between these two quantities, and determines, possibly with an error, the larger of the two. The comparison in this second step is well described by a random walk of decision.

FACTORS AFFECTING THE DECISION PROCESS

In the standard random walk model, the barrier that the process has to hit is fixed. Suppose now that the information available to the decision maker in two tasks is different, and is of better quality in one of them.

For example, in a risky choice the DM has a precise statement on the probability of the outcomes in the lotteries she has to choose from. In the ambiguous choice, on the contrary, she has only limited information on these probabilities. She must provide an estimate on the likelihood of different outcomes on the basis of some reasonable inference. Similarly, in a choice of lotteries that are paid at different points in time, lotteries paid in the current period are easier to analyze than those paid further in, say, 1 month, because the decision maker has to consider which different contingencies may occur in the next month, and how they might affect the outcome and the utility for her of different consequences. Consider now the prediction of this model on the response time and error rate in the two cases. Intuitively, a harder task should take longer. This is what the random walk model predicts: if the distance from the initial point that the process has to cover is the same, and the process is slower when the information is worse, then the response time should be longer in the harder process. However, we observe the opposite: the response time in the ambiguous choice is consistently shorter than in the risky choice.

A consideration of the extreme case in which the signal that is observed is completely non-informative reveals what might be the missing step. Suppose that indeed the signal provides no information. In this case, waiting to observe the signal provides no improvement over the immediate decision. Since waiting typically implies a cost (at least an opportunity cost of time that could be better used in other ways), the decision in this case should be immediate, because delay only produces a waste of time. So, in the case of the worse possible signal, the response time is the shortest. This conclusion seems to contradict the prediction of the random walk, but instead it contradicts only the assumption that the barrier the process has to hit is fixed. The distance from the initial point at which the process stops should instead depend on the quality of the signal: everything else being equal, a better signal is worth being observed for a longer time.

In the next section we make this informal argument more formal, by showing precisely that when the quality of the signal is better, two opposing factors are active: first, the quality of the signal advises to wait and get better information. This counteracts the second, direct effect (proceeding with a better signal is faster), and may produce what we observe: longer response times with the better, more informative signal.

A Simple Example

The intuitive content of the model can be appreciated better if we consider first the very simple decision problem already introduced in the Formal model section above. If the decision maker receives no additional information, the value for her problem is

$$v(p) = \max\{p, 1 - p\}$$

with the optimal choice of r if $p \leq 0.5$, and l otherwise.

Suppose now that the decision maker can observe instances of an informative signal on the state: the function from the true state to a signal space is called, using a term of statistical theory, an experiment. She can observe the signal produced by the experiment for as many periods as she wants, but the final utility will then be discounted by a factor δ. Now, it is no longer necessarily optimal to choose immediately on the basis of the prior belief; rather, it may be better to wait, observe the signal, update the belief, and make a better choice. Since the value of the reward is discounted, the decision maker has a genuine problem: she has to decide between collecting information, and choosing immediately.

Assume for the moment that an optimal policy, for a given initial belief p, exists. The value of the problem computed at the optimal policy for any such initial belief defines the value function for the problem, which we denote by V. This function is obviously larger than v, since the decision maker has the option of stopping immediately. It is known that the optimal policy for this decision maker can be described as a function of the belief she has regarding the state – that is, on the current value of p. The way in which this dependence works is clear. For a belief p at which $V(p) = v(p)$, the optimal policy is to do what yields $v(p)$; namely, to stop. For the values for which $V(p) > v(p)$, since stopping would only give $v(p)$, the decision maker has to continue experimenting. It turns out in this simple example that there is a cutoff belief, call it p^*, such that it is optimal to stop if, and only if, $p \geq p^*$ or (symmetrically) $p \leq 1 - p^*$.

Consider now the effect on the decision to stop when the quality of the signal provided to the decision maker improves. Introducing a notation used later, we denote the experiments P and Q, with P more informative than Q. Note that the function v does not depend on the experiment, but the value function V and the cutoff belief p^* depend on it, and we write, for example, $V(P, \cdot)$ and $p^*(P)$ to make this dependence explicit. When P replaces Q, the value function V becomes larger, because the information is better (this is intuitively clear, and is proved formally below). Therefore, the set of beliefs at which V is equal to v becomes smaller; that is, the critical belief p^* becomes larger: $p^*(P) \geq p^*(Q)$. Note for future reference that this value also depends on the other parameters of the problem, in particular the discount factor δ, although we do not make this dependence explicit in the notation.

Quality of the Signal and Response Time

What is the effect of this change on the response time? An increase in the value of p^* tends, everything else being equal, to make the response time longer: it takes more observation to reach a cutoff which is farther from the initial belief. Since an improvement of the signal increases p^*, this direct effect would by itself produce a longer response time. However, a better signal also reduces the time needed to reach a fixed cutoff belief, since the information is more effective.

The net effect is studied below for a more general class of problems, but it is easy to see intuitively what it is. Consider first the case in which the signal provides no information at all. In this case there is no point in waiting and experimenting, and therefore the optimal policy is to stop immediately. Consider now the case in which the experiment provides complete information: as soon as the signal is observed, the state is known for sure. In this case, the optimal waiting time is at most one period: if the decision maker decides to experiment at all, then she will not do it for longer than one period, since in that single period she gets all the information she needs, and additional signals are useless. Note that these two conclusions are completely independent of the value of the discount, since our argument has never considered this value.

Consider now the case of an experiment of intermediate quality between the two extremes just considered: the experiment provides some information, so the posterior belief is more accurate, but the information is never enough to reach complete certainty. If the discount factor becomes closer to 1, then the opportunity cost of gathering additional information becomes smaller. The value of a utility at T is scaled down by a factor δ^T, which is close enough to 1. So if we keep the information fixed, and consider larger and larger values of δ, we see that the cutoff belief p^* increases. Since the experiment is fixed, the effect on the time to reach this cutoff now is unambiguous. Note that in fact the time to stop is a function of the history of signals observed. The probability distribution on this set is given by the experiment. Since the cutoff is higher, for any history the time to reach this cutoff increases, and it is easy to see that we can make it arbitrarily large.

We can now conclude that the time to decide (the response time that we observe) is a hill-shaped function of the quality of information. This conclusion holds in a more general model, which is presented in the Appendix to this chapter.

COGNITIVE ABILITIES AND PREFERENCES

We present how the model we have developed so far can explain experimental as well as real-life choice behavior of a large group of subjects, relating the choice made in different environments to cognitive abilities. Economic theory makes no statement regarding the correlation between characteristics of individual preferences in different domains. For example, the coefficient of risk aversion is considered independent of the impatience parameter. Also, no correlation is assumed between these preferences and the cognitive ability (CA) of the individual. The predictions of the theory of choice that we have presented are different.

How can cognitive abilities affect preferences? In the theory we have developed so far, the utility of an option is perceived with a noise. The more complex the option is, the larger the noise in the perception. For example, evaluating the utility of a monetary amount paid for sure is easy, and no one has any doubts when choosing between \$10 and \$15. Instead, evaluating a lottery giving on average \$10 is harder, and it is harder still to compare the choice between two lotteries. Similarly, the utility of \$15 to be paid on 10 days is not as sharply perceived as the same payment immediately: we have to consider several different possible intervening factors, such as the impossibility of getting or receiving the payment, other payments that can be received in the same interval, and so on.

Different degrees of CA make the perception of an option more or less sharp. Consider now the choice between a certain amount and a lottery. While the utility of the first is perceived with precision by every individual, the noise around the second one increases for individuals with a lower CA, and so that option is less likely to be chosen by those individuals: subjects with a lower CA make more risk-averse choices. Similarly, in the choice between a payment now and one in the future, they perceive the second more noisily than the first, and so they are less likely to choose it, and they make more impatient choices. The theory predicts that impatience and risk aversion are correlated, and these in turn are correlated with cognitive abilities.

Test of the Theory

We examined whether and how attitudes to risk, ambiguity, and inter-temporal choices are related in a large ($N = 1066$) sample of drivers in an important national (USA) company (see Burks *et al.*, 2007). Thanks to an agreement with the company, we ran extensive (4 hours) laboratory experimental testing with the participating subjects on a battery of tasks involving choice under risk, ambiguity, choice over time delayed payments, as well as a variety of psychological measurements and cognitive tasks (see Burks *et al.*, 1943 for a detailed description of the experiment). Similar results, which confirm the robustness of ours, can be found in Benjamin *et al.* (2007); Dohman *et al.* (2007). From a different perspective, the issue of the connection between cognitive abilities (specifically numeracy) and decision making can be found in Peters *et al.* (2006).

We had three separate measures of CAs: a measure of the IQ (Raven's matrices), a measure of numerical ability (Numeracy) on tests provided by the ETS (Educational Testing Service), and the score on a simple game played against the computer (called *Hit 15*, because the game is a race between two players to reach position 15 on a gameboard) which measures the planning ability of the individual.

In the choice under uncertainty, subjects were asked to choose between a fixed lottery and a varying certain amount. The lottery was either risky (with known, equal probability of the two outcomes) or ambiguous (unknown probability of two colors, and the subject was free to pick the winning color). In choices of different profiles of payments, subjects had to choose between two different payments at two different points in time, a smaller payment being paid sooner.

A first clear effect due to CA was the number of errors the subject made, if we define error (as before) as the number of switches between certain amount and lottery above two. We found that inconsistency increases with our measures of CA, in particular IQ and *Hit 15* score.

The effect of CA on preferences was as predicted: the patience and the index of cognitive ability are positively correlated. Also, risk aversion and the index of cognitive ability are negatively correlated. As a result, there is a negative correlation between risk aversion and impatience.

The effect of the difference in cognitive ability extends to behavior in strategic environments. In our experiment, subjects played a discrete version of the trust game: both players were endowed with $5; the first mover could transfer either $0 or the entire amount, and the second player could return any amount between $0 and $5. Both amounts were doubled by the experimenter. Before the choice, subjects reported their belief on the average transfer of the participants in the experiment both as first and as second movers.

We found that a higher IQ score makes a subject a better predictor of the choice of the others as first movers: while the average underestimates the fraction of subjects making a $5 transfer, subjects with higher IQ are closer to the true value. Similarly, they are closer to the true value of the transfers of second movers. The behavior is also different. As second movers, subjects with higher IQ make higher transfers when they have received $5, and smaller transfers in the opposite case.

The behavior as first movers is more subtle to analyze, since beliefs also enter into the choice: since subjects with higher IQ believe that a larger fraction of second movers will return money, they might be influenced by this very fact. In addition, the difference in risk aversion might affect choices. Once we control for these factors, however, subjects with higher IQ are more likely to make the $5 transfer.

We also followed the performance on the workplace in the months following the initial collection of experimental data; in particular, the length of time the subject remained with the company, and, when relevant, the reason for quitting the job. In the training offered by the company, quitting before a year can be safely considered to be evidence of poor planning: trainees leave the company with a large debt (for the training costs have to be paid back to the company if an employee quits before the end of the first year), they have earned little, and have acquired no useful experience or references for their resumé. If we estimate the survival rate for different socio-economic variables (for example, the married status), then the variables have no significant effect on the survival rate, while the *Hit 15* affects it largely and significantly.

APPENDIX: RANDOM WALK WITH ENDOGENOUS BARRIERS

We denote the unknown parameter (for example, the state of nature) as $\theta \in \Theta$. The decision maker has an initial belief on the parameter, $\mu_0 \in \Delta(\Theta)$, and has to take an action $a \in A$. The utility she receives depends on the state of nature and the action taken, and is described by a function $u : \Theta \times A \rightarrow R$. She can, before she takes the action, observe a signal $x \in X$ with a probability that depends on the state of nature, denoted for example by $P_\theta \in \Delta(X)$. In classical statistical terminology, this is an experiment $P \equiv (P_\theta)_{\theta \in \Theta}$. For any given prior belief on the set Θ, this experiment induces a probability distribution on the set of signals:

$$P_\mu(x) = \int_\Theta \mu(d\theta) P_\theta(x).$$

The subject can observe independent replications of the signal as many times as she likes, stop, and then choose an action $a \in A$. The use of the experiment has a fixed cost c for every period in which it is used.

The information she has at time $t \in \{0, 1, \ldots\}$ is the history of signals she has observed, an element

$(x_0, \ldots, x_{t-1}) \in X^t$. The posterior belief at any time t is a random variable dependent on the history of signals she has observed, and is denoted by μ_t. Let $B(\mu, x)$ denote the posterior belief of a Bayesian decision maker with a prior belief μ after observing a signal x. We write $B(\mu, x, P)$ if we want to emphasize the dependence of the updating function on the experiment P.

The decision maker can make two separate choices in each period: first, whether to stop observing the signal, and second, if she decides to stop, which element of A to select. The action she chooses at the time in which she stops is optimal for her belief at that time. If her posterior is ν, her value at that time is equal to $v(\nu)$, the expected value conditional on the choice of the optimal action, namely:

$$v(\nu) = \max_{a \in A} E_\nu u(\cdot, a) \tag{4.7}$$

Conditional on stopping, the action in A is determined by the maximization problem we have just defined, and the value of stopping is given by v. We can therefore focus on the choice of when to stop.

A policy π is a sequence of mappings $(\pi_0, \ldots, \pi_t, \ldots)$, where each π_t maps the history of observations at time t, $(x_0, \ldots, x_{t-1})$ into {0,1}, where 1 corresponds to *Stop*. The first component π_0 is defined on the empty history.

The initial belief μ and the policy π define a probability distribution over the set of infinite histories X^∞, endowed with the measurable structure induced by the signal. We denote by $E_{\pi,\mu}$ the corresponding expectation. Also there is a stopping time T (a random variable) determined by the policy π, defined by

$$T \equiv \min_t \{\pi_t(x_0, \ldots, x_{t-1}) = 1\}$$

The expected value at time zero with the optimal policy depends on the signal the subject has available, and is given by

$$V(\mu, P) = \max_\pi E_{\pi,\mu} \left[\delta^T v(\mu_T) - \sum_{t=1}^{T} c\delta^{t-1} \right]$$

where we adopt the convention that $\sum_{t=1}^{0} c\delta^{t-1} = 0$.

Normally distributed signals An important example is the class of normally distributed experiments. Let Θ be a subset of the real line, indicating the expectation of a random variable.

An experiment P is defined as the observation of the random variable X distributed as $N(\theta, \sigma^2)$, where the variance σ^2 is known. An experiment Q, given by the observation of the variable Y distributed as $N(\theta, \rho^2)$, is dominated by P if, and only if, $\rho > \sigma$. This is in turn equivalent to the existence of a normal random variable Z with zero expectation and variance equal to $\rho^2 - \sigma^2$ such that

$$Y = X + Z.$$

OPTIMAL POLICY

The operator M on the space of continuous functions on $\Delta(\Theta)$ with the sup norm is defined by

$$M(P, W)(\mu) \equiv \max\{v(\mu), -c + \delta E_{\mu,P} W(B(\mu, \cdot))\}$$

where the function v is defined in (4.7).

This operator is a contraction on that space, because it satisfies the conditions of Blackwell's theorem. Hence the value function V exists, and is the solution of the functional equation:

$$V(\mu, P) = M(P, V)(\mu), \text{ for every } \mu.$$

The value function equation describes implicitly the optimal policy, which is a function $\hat{\Pi}$ of the current belief. As in our simple example, the policy is to stop at those beliefs in which the value function V is equal to the value of stopping immediately, v. Formally we define the stopping time region $S(P) \subseteq \Delta(\Theta)$ as

$$S(P) \equiv \{\mu : v(\mu) = V(\mu, P)\}$$

The optimal policy is stationary: the function π_t depends on the history of signals only through the summary given by the current belief. This optimal policy is described by the function $\hat{\Pi}$

$$\hat{\Pi}(\mu) = 1 \quad \text{if and only if } \mu \in S(P).$$

VALUE AND QUALITY OF SIGNALS

Consider now two experiments of different quality, P and Q say. Let $\succeq$ denote the partial order (as defined by Blackwell, 1951; Targerson, 1991) over experiments. We now show that if the experiment is more informative, then the set of beliefs at which the decision maker continues to observe the signal is larger than it is for the worse signal.

Theorem 4

1. *The operator M is monotonic in the order $\succeq$, that is, for every function $W : \Delta(\Theta) \to R$, if $P \succeq Q$, then $M(P, W) \geq M(Q, W)$, and therefore $V(\cdot,P) \geq V(\cdot,Q)$*
2. *The optimal stopping time region S is monotonically decreasing, namely if $P \succeq Q$ then $S(P) \subseteq S(Q)$.*

In terms of our main application, decision under risk and uncertainty, the conclusion is that with a richer information (risk) the barrier where the random walk stops is farther than it is with the more poor information (ambiguity). As a consequence, the updating process may take longer in risk than in ambiguous choices.

Quality of signals and response time We now present formally the argument presented informally in our analysis of the simple example. Recall first that:

1. An experiment is called totally un-informative, denoted by P^u, if

$$\text{for all } \theta^1, \theta^2 \in \Theta, P^u_{\theta^1} = P^u_{\theta^2}$$

2. An experiment is called totally informative, denoted by P^u, if

$$\text{for all } \theta^1, \theta^2 \in \Theta, P^u_{\theta^1} \perp P^u_{\theta^2},$$

that is the two measures are mutually singular.

We now have:

Lemma 5

1. *If the experiment P is totally informative, then at the optimal policy the stopping time $T \leq 1$, $(\pi, \mu) - a.s.$;*
2. *If the experiment P is totally un-informative, then at the optimal policy the stopping time $T = 0$, $(\pi, \mu) - a.s.$;*

As in the analysis of our simple example, note that the two conclusions are independent of the discount factor δ and the cost c.

We now turn to the analysis of the response times when the experiments have intermediates, namely for experiments P such that $P^i \succeq P \succeq P^u$.

Define the function

$$U(\mu) \equiv \int_\Theta \max_{a \in A} u(\theta, a) d\mu(\theta)$$

This is the value at the belief μ of a decision maker who is going to be completely and freely informed about the state before she chooses the action. Information is always useful for her (for every belief that is different from complete certainty about a state) if the value of the optimal choice at μ is smaller than the expected value when complete information will be provided:

Assumption 6

Information is always useful, namely

$$\text{For every } \mu \in \Delta(\Theta), \text{ if } \mu \notin \{\delta_\theta : \theta \in \Theta\}, \text{ then } U(\mu) > v(\mu). \tag{4.8}$$

Assumption 7

An experiment P is intermediate, that is:

1. *For every finite number n of independent observations, and initial belief in the relative interior of $\Delta(\Theta)$,*

$$B(\mu, P^n)$$

is in the relative interior of $\Delta(\Theta)$.

2. *As the number of independent observations tends to infinity, the product experiment converges to the totally informative experiment.*

Theorem 8

If information is always useful (assumption 9.3) and the experiment is intermediate (assumption 9.4), then

$$\lim_{c \downarrow 0, \delta \uparrow 1} T = \infty, (\pi, \mu_0) - a.s.$$

where π is the optimal policy.

References

Anscombe, A. and Aumann, R. (1963). A definition of subjective probability. *Ann. Math. Stat.* 34, 199–205.

Benjamin, D., Brown, S., and Shapiro, J. (2007). *Who is "behavioral"*? Discussion Paper.

Blackwell, D. (1951). Comparison of experiments. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, CA: University of California, 93–102.

Burks, S., Carpenter, J., Goette, L. *et al*. (1943). Using behavior economic experiments at a large motor carrier: the context and design of the truckers and turnover project. IZA DP No 2789.

Burks, S., Carpenter, J., Goette, L., and Rustichini, A. (2007). Cognitive abilities explain economic preferences, strategic behavior and job performance. Working Paper, University of Minnesota.

Cartwright, D. (1941a). The relation of the decision time to the categories of response. *Am. J. Psychol.* 54, 174–196.

Cartwright, D. (1941b). The decision time in relation to the differentiation of the phenomenal field. *Psychological Rev.* 48, 425–442.

Cartwright, D. and Festinger, L. (1943). A quantitative theory of decision. *Psychological Rev.* 50, 595–621.

Davidson, D. and Marschak, J. (1959). Experimental tests of stochastic decision theory. In: C.W. Churchman (ed.), *Measurement Definitions and Theories*. New York, NY: John Wiley & Sons, pp. 233–269.

Debreu, G. (1958). Stochastic choice and cardinal utility. *Econometrica* 26, 440–444.

Dickhaut, J., McCabe, K., Nagode, J. *et al*. (2003). The impact of the certainty context on the process of choice. *Proc. Nat. Acad. Sci.* 100, 3536–3541.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2007). Are risk aversion and impatience related to cognitive abilities. IZA DP 2735.

Gul, F. and Pesendorfer, W. (2003). Random expected utility. *Econometrica* 74, 121–146.

Krantz, D., Luce, R., Suppes, P., and Tversky, A. (1971). *Foundations of Measurement,* vol. I, *Additive and Polynomial Representations*. New York, NY: Academic Press.

Köbberling, V. (2006). Strength of preferences and cardinal utility. *Economic Theory* 27, 375–391.

Maccheroni, F., Marinacci, M., and Rustichini, A. (2007). Preference based decision process. Working Paper, University of Minnesota.

McFadden, D. and Richter, M. (1991). Revealed stochastic preferences. In: J.S. Chipman, D. McFadden, and M.K. Richter (eds), *Preferences, Uncertainty and Optimality*. Boulder, CO: Westview Press, pp. 161–186.

Moyer, R. and Landauer, T. (1967). Time required for judgements of numerical inequality. *Nature* 215, 1519–1520.

Padoa-Schioppa, C. and Assad, J. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223.

Peters, E., Vastfjall, D., Slovic, P. *et al*. (2006). Numeracy and decision making. *Psychological Sci.* 17, 407–413.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Rev.* 85, 59–108.

Rustichini, A., Dickhaut, J., Ghirardato, P. *et al*. (2005). A brain imaging study of the choice procedure. *Games Econ. Behav.* 52, 257–282.

Savage, L.J. (1954). *The Foundation of Statistics*. New York, NY: Dover Publications Inc.

Schall, J.D. (2001). Neural basis of deciding, choosing and acting. *Nat. Rev. Neurosci.* 2, 33–42.

Shadlen, M.N. and Newsome, W.T. (1996). Motion perception: seeing and deciding. *Proc. Natl Acad. Sci. USA* 93, 628–633.

Shadlen, M.N. and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (Area LIP) of the rhesus monkey. *J. Neurophysiol.* 86, 1916–1936.

Shapley, L. (1975). *Cardinal utility comparisons from intensity comparisons*. Report R-1683-PR, The Rand Corporation, Santa Monica, CA.

Sigman, M. and Dehaene, S. (2005). Parsing a cognitive task: a characterization of mind's bottleneck. *PLOS Biol.* 3, e37.

Smith, P. (2000). Stochastic dynamic models of response time and accuracy: a foundational primer. *J. Math. Psychol.* 44, 408–463.

Torgersen, E. (1991). *Comparison of Statistical Experiments*. Cambridge: Cambridge University Press.

von Neumann, J. and Morgenstern, O. (1947). *Theory of Games and Economic Behavior*, 2nd edn. Princeton, NJ: Princeton University Press.

C H A P T E R

5

Experimental Neuroeconomics and Non-cooperative Games

Daniel Houser and Kevin McCabe

O U T L I N E

INTRODUCTION

Embodied brain activity leads to emergent computations that determine individual decisions. In turn, individual decisions, in the form of messages sent to an institution, lead to emergent computations that determine group-level outcomes. Computations can be understood in terms of a set of transformation rules, the encoding of information, and the initial conditions that together produce the computation, and we will refer to these three elements together as a *computational mechanism* or simply a *mechanism*. Neuroeconomics is interested in understanding the interrelationship between those mechanisms that exist in our evolved brains and those that exist in our constructed institutions, and their joint computation.

Game theory provides a nice middle ground for neuroeconomics studies, because it links individual decision making to group-level outcomes with a clearly defined mechanism. The mechanism is the game tree, which specifies who gets to move when, what moves they can make, what information they have when they make their move, and how moves of different players interact to determine a joint outcome over which the players have varied interests. Non-cooperative game theory has played an important role in economic thinking, starting with the studies of imperfect competition in the late 1800s, but it was the publication of von Neumann and Morgenstern's (1947) book, followed shortly by John Nash's (1950) formulation of and proof of existence of Nash equilibrium that gave game theory its modern form. In 1994, the Nobel Prize in Economic Sciences was awarded to John Harsanyi (1920–2000), John Nash (1928–), and Reinhard Selten (1930–).

As game theory has grown in popularity, many books have become available to the reader. In addition to our review below, and the reference therein to original works, an accessible treatise is Osborne (2004).

Extensive Form Games

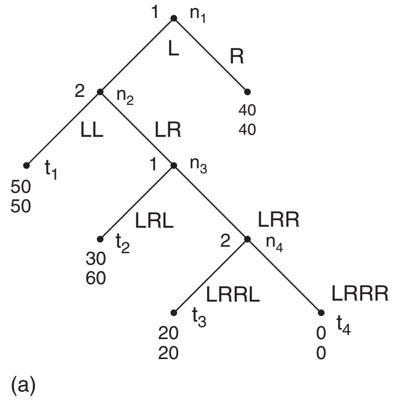
A game involves two or more players. Figure 5.1 depicts two-person games. Figure 5.1a is an example of an extensive form game with perfect information (Kuhn, 1950). The game consists of nodes and branches that connect the nodes, called the game tree. The nodes n_1–n_4 are called *decision nodes*, as they each have branches connecting them to later nodes, and the nodes t_1–t_5 are called *terminal nodes*. Each terminal node has a payoff vector associated with it where the top number is decision maker 1's payoff and the bottom number is decision maker 2's payoff. For convenience the branches have been labeled, L, R, LL, LR, LRL, etc. To the top left of each decision node is a number, 1 or 2, indicating that the decision maker owns, or equivalently gets to move at, that node. Decision maker 1 owns n_1 and n_3.

Pure Strategy Nash Equilibrium

A play of the game is a connected path through the game tree that starts at n_1 and ends at a terminal node. Thus, (L, LR, LRL) is a play that ends at t_3. A pure strategy for a player is a choice of branch for each decision maker at each decision node that he owns. For decision maker 1, let X = {(L, LRL), (L, LRR), (R, LRL), (R, LRR)} denote the set of all possible pure strategies, and let $x \in X$ be a particular pure strategy. Similarly, for decision maker 2, let Y = {(LL, LRRL), (LL, LRRR), (LR, LRRL), (LR, LRRR)} denote the set of all pure strategies, and let $y \in Y$ be a particular strategy. Each strategy pair (x, y) determines a play through the game tree. Thus $x' = (\text{L, LRL})$ and $y' = (\text{LR, LRRL})$ determine the play (L, LR, LRL), as does the strategy pair $x'' = (\text{L, LRL})$ and $y'' = (\text{LR, LRRR})$. The payoffs for decision makers 1 and 2 can be denoted $P(x, y)$ and $Q(x, y)$, respectively. For example, $P(x', y') = 30$ and $Q(x', y') = 60$.

A Nash Equilibrium of a game is a strategy pair (x^*, y^*) such that the following conditions hold:

$$P(x^*, y^*) \geq P(x, y^*) \text{ for all } x \in X \tag{5.1}$$



DM1

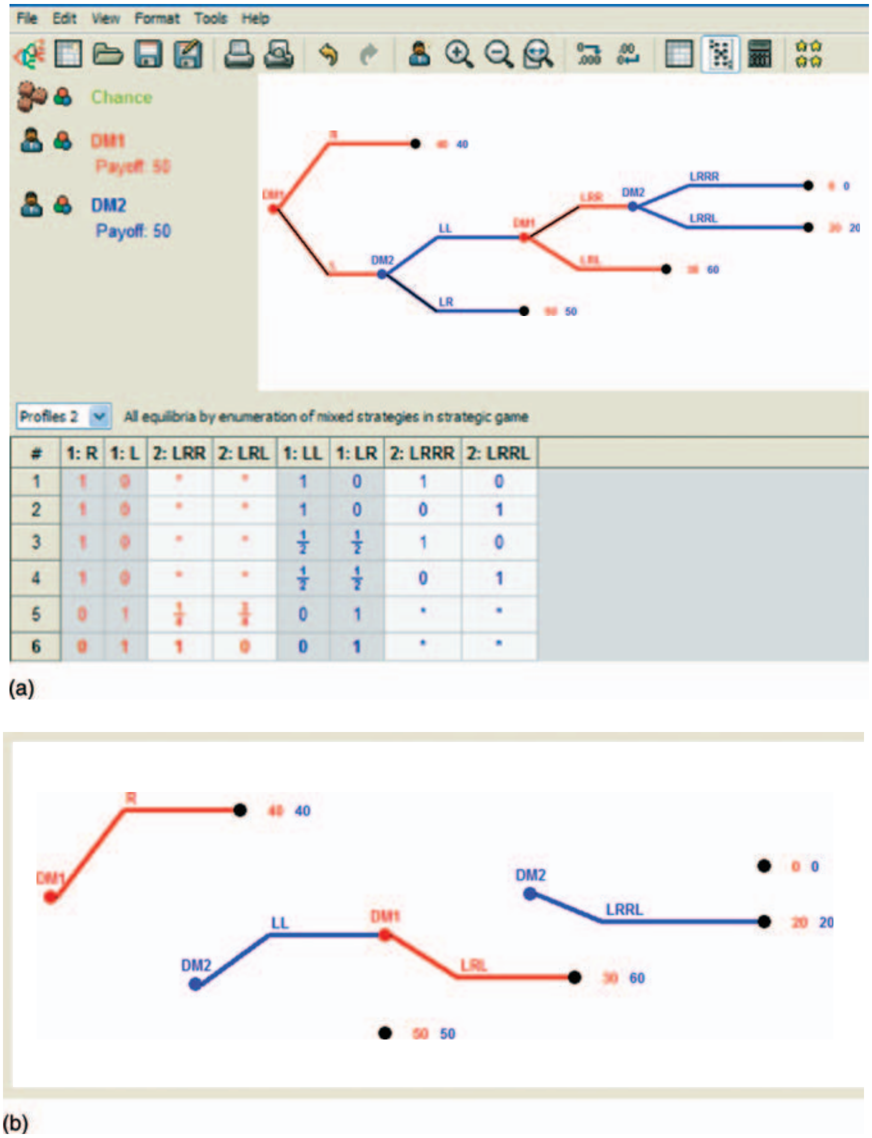
DM2	L, LRL	L, LRR	R, LRL	R, LRR
LL LRRL	50 50	50 NE 50	40 40	40 40
LL LRRR	50 50	50 NE 50	40 40	40 40
LR LRRL	30 60	20 20	40 NE 40	40 NE 40
LR LRRR	30 60	0 0	40 NE 40	40 NE 40

(b)

FIGURE 5.1 Simple example of a two-person game: (a) game in extensive form; (b) game in normal form. Figure 5.1(a) shows a finite extensive form game with perfect information. Play starts when decision maker 1 (DM1) moves left (L) or right (R) at the decision node n_1. If DM1 moves right, then the game ends at decision node t_0 and each player gets a payoff of 40. Note the top payoff goes to DM1 and the bottom payoff goes to DM2. A pure strategy for each decision maker is a choice of move at every decision node he controls. Thus, a strategy may be (L, LRL) for DM1 and (LL, LRRL) for DM2. Notice a strategy must specify what DM2 will do at decision node n_4, even thought LL terminates the game at t_1. Every pair of strategies produces a play of the game. For example the strategies given above produce the play (L, LL, t_1). Game theory assumes that players behave rationally in choosing strategies that maximize their expected payoff from the resulting play of the game. Figure 5.1(b) shows the game in Figure 5.1(a) in strategic or normal form. Notice now we list all the strategies that decision makers have and then in the cell indicate the payoff that results from the play of the game. The pure strategy Nash Equilibrium (NE) of a game is simply a choice of strategies by DM1 and DM2 such that neither decision maker could do better by unilaterally changing only his strategy. The pure strategy Nash Equilibria of the strategic game are noted. These are also the Nash Equilibria of the game in Figure 5.1(a).

$$Q(x^*, y^*) \geq Q(x^*, y) \text{ for all } y \in Y \tag{5.2}$$

From the definition, it is clear that a candidate strategy pair (x', y') for a Nash Equilibrium can be rejected if we can find a strategy for either player that leads to a better outcome for that player given the



All equilibria by enumeration of mixed strategies in strategic game

#	1: R	1: L	2: LRR	2: LRL	1: LL	1: LR	2: LRRR	2: LRRL
1	1	0	*	*	1	0	1	0
2	1	0	*	*	1	0	0	1
3	1	0	*	*	$\frac{1}{2}$	$\frac{1}{2}$	1	0
4	1	0	*	*	$\frac{1}{2}$	$\frac{1}{2}$	0	1
5	0	1	$\frac{1}{4}$	$\frac{3}{4}$	0	1	*	*
6	0	1	1	0	0	1	*	*

FIGURE 5.2 Solving the game using *Gambit*: (a) all Nash Equilibria of the game; (b) subgame perfect Nash Equilibrium of the game. *Gambit* is a software tool that allows you, with care, to enumerate all of the Nash Equilibria of a finite game. In Figure 5.2(a), we see the game in Figure 5.1(a) depicted in *Gambit*. Below the tree we see all of the Nash Equilibria of the game including mixed strategies. *Gambit* also has features that allow us to solve for the subgame perfect equilibrium. This is shown in Figure 5.2(b), which shows that the strategy (R, LRL) and the strategy (LL, LRRL) is a subgame perfect NE. Note some of the branch labels have been changed. *Gambit* also allows us to solve for the quantal response equilibrium (QRE) of the game described later in this chapter.

other player's strategy – i.e., if either inequality below is true:

$$P(x, y') > P(x', y') \text{ for some } x \in X, \text{ or}$$

$$Q(x', y) \geq Q(x'', y') \text{ for some } y \in Y.$$

Thus a Nash Equilibrium strategy pair is a pair that cannot be rejected. If the inequalities in equations (5.1) and (5.2) are replaced with strict inequality signs, then we call the pair (x^*, y^*) a *Strict Nash Equilibrium*.

For example, $x^* = (\text{L, LRR})$ and $y^* = (\text{LL, LRRL})$ is a Nash Equilibrium in our game above. For a more general game with m players, Nash Equilibrium (see Nash, 1950) is defined as above only with m simultaneous inequalities. On the other hand, $x' = (\text{R, LRR})$ and $y' = (\text{LL, LRRL})$ is not a Nash Equilibrium of the game since $P(x^*, y') > P(x', y')$.

A number of attempts have been made to write software that can calculate all of the Nash Equilibria of a game. One such example is *Gambit*, co-authored by Richard McKelvey, Andrew McLennan, and Theodore Turocy (2007), which can be downloaded at http://gambit.sourceforge.net. Sample *Gambit* output for the game above is shown in Figure 5.2a. *Gambit* found six Nash Equilibria, including three that involve *mixed strategies* (described later). The fact that a game can have more than one Nash Equilibrium has inspired many attempts to define refinements of the Nash Equilibrium.

Subgame Perfect Equilibrium

One important refinement of Nash Equilibrium, due to Reinhard Selten (1975), is known as the subgame perfect equilibrium, or simply SPE, of an extensive form game. A feature of an SPE is that players have no incentive to deviate from it during the course of the game. That is, an SPE is a Nash Equilibrium strategy no matter where in the game the player begins.

To see this more clearly, consider our example game in Figure 5.1a. Note that each of the decision nodes n_2–n_4 describes a subgame of the original game by simply ignoring what went before the particular node. Treat each of these nodes as a starting node of the new subgame. A strategy profile (x^*, y^*) is a SPE if the relevant parts of the profile are also a Nash Equilibrium for each subgame of the original game. So, for example, while x = (L, LRR) and y = (LL, LRRL) is a Nash Equilibrium of the game, it is not an SPE since [(*, LRR), (*, LRRL)] is not a Nash Equilibrium for the subgame starting at n_3 – i.e., decision maker 1 would strictly prefer to play LRL. If we define the length of a subgame as the longest path to a terminal node, then we can find all of the subgame perfect equilibria of a game by working backwards through the game tree and finding all of the Nash Equilibrium of the subgames, starting with subgames of the shortest length, the next shortest length, etc. For our example, y^* = (*, LRRL) is the Nash Equilibrium of the subgame starting at n_4, x^* = (*, LRL); y^* = (*, LRRL) is the Nash Equilibrium of the subgame starting at n_3, x^* = (*, LRL); y^* = (LL, LRRL) is a Nash Equilibrium of the subgame starting at n_2; and, finally, x^* = (R, LRL), y^* = (LL, LRRL) is a Nash Equilibrium of the game starting at n_1. These calculations are also shown in the *Gambit* output illustrated in Figure 5.2b. Kuhn (1953) proved that every finite extensive form game with perfect information has a SPE.

Normal or Strategic Form Games

The extensive form game shown in Figure 5.1a has an equivalent strategic or normal form, as shown in Figure 5.1b. In a strategic form game, each player has to make a choice from a set of choices. Player DM1 chooses one of the four columns, where each column represents a pure strategy choice. Simultaneously, player DM2 chooses one of four rows corresponding to one of DM2's pure strategies. Players' choices together select one of the 16 cells in the 4 × 4 matrix game. The cell selected is the outcome from playing the game, and each cell contains the payoff to the players if that cell is selected.

The set of Nash Equilibria for the normal form of this game is exactly identical to that described above for this game's extensive form. Indeed, the set of Nash Equilibria for any given strategic form game is the same as the set of Nash Equilibria for that same game expressed in extensive form. The reason is that Nash Equilibrium is defined in terms of available strategies. It does not matter to Nash Equilibrium analysis when those strategies are executed, or how they are described.

Mixed Strategy Equilibrium

A difficulty with our definition of Nash Equilibrium described above is that not every game has such a Nash Equilibrium. The difficulty can be seen in the Rock–Scissors–Paper example shown in Figure 5.3a. In this game, both decision makers must simultaneously choose Rock (R), Scissors (S) or Paper (P). When their choices are revealed, their payoffs are as follows: If they both choose the same, then they each get zero. If P and R are played, then P wins and the loser must pay the winner $1. If S and P are played, then S wins and the loser must pay the winner $1. Finally, if R and S

		DM2		
		R	S	P
DM1	R	0 0	−1 +1	+1 −1
	S	+1 −1	0 0	−1 +1
	P	−1 +1	+1 −1	0 0

(a)

		DM2	
		R	S
DM1	R	0 0	−1 +1
	S	+1 −1	0 0
	P	−1 +1	+1 −1

(b)

FIGURE 5.3 Mixed strategy Nash Equilibrium: (a) Rock–Scissors–Paper; (b) DM2 cannot play Paper. Figure 5.3(a) depicts a strategic game that has only a mixed strategy equilibrium of (1/3, 1/3, 1/3) for each decision maker and an expected payoff of 0. The game depicts the zero sum game Rock–Scissors–Paper. Figure 5.3(b) shows a truncated Rock–Scissors–Paper game where DM2 is not allowed to play Paper (P). Since Paper beats Rock and DM2 can't play Paper, this should incline DM1 to play Rock more often, but not all the time since this would be predictable and lead only to (R, R). Instead, DM1 will play Rock 2/3 of the time, leading DM2 to play Rock 2/3 of the time, leading to a payoff of 4/9 for DM1.

are played, then R wins and the loser must pay the winner $1. To see that there is no Nash Equilibrium as defined above, note that if DM1 plays R then DM2 strictly prefers P, but if DM2 plays P then DM1 strictly prefers S, and so on, resulting in an endless cycle, or equivalently no solution to the inequalities (5.1) and (5.2) above. However, we can find a Nash Equilibrium of this game if we allow DM1 and DM2 to choose mixed strategies as defined below.

We can denote a mixed strategy for decision makers 1 and 2 as probability distributions, p and q, over X and Y respectively. Thus, for example, the probability that DM1 plays a particular strategy x is $p(x)$. For convenience, we can assume the players' preferences regarding strategies are ordered according to expected payoffs (Bernoulli, 1738; von Neumann and Morgenstern, 1944; see also Chapter 3 of this volume). Thus, for any given p and q, DM1 and DM2's respective payoffs are given by:

$$\mathrm{EP}(p, q) = \sum_{x \in \mathrm{X}} \sum_{x \in \mathrm{Y}} p(x) q(y) \mathrm{P}(x, y)$$

$$\mathrm{EP}(p, q) = \sum_{x \in \mathrm{X}} \sum_{x \in \mathrm{Y}} p(x) q(y) \mathrm{Q}(x, y)$$

A pure strategy (x, y) is now a special case of a mixed strategy where $p(x) = 1$ and $q(y) = 1$. A mixed strategy Nash Equilibrium is a p^*, q^* such that $\mathrm{EP}(p^*, q^*) \geq \mathrm{EP}(p, q^*)$ for all mixed strategies p, and $\mathrm{EQ}\ (p^*, q^*) \geq \mathrm{EQ}\ (p^*, q)$ for all mixed strategies q. For the Rock–Scissors–Paper game, there is one Nash Equilibrium in mixed strategies $p = (1/3,\ 1/3,\ 1/3)$ and $q = (1/3, 1/3, 1/3)$.

More generally, if we have a strategic form game with n players, indexed by i, each of whom have von Neumann-Morgenstern preferences, then we can define player i's mixed strategy as p_i and we can define all the remaining strategies of the $n - 1$ players as $p_{-i} = (p_1, \ldots, p_{i-1}, p_{i+1}, \ldots, p_n)$ and payoffs can be defined by $\mathrm{EU}_i(p_i, p_{-i})$. We can now extend our definition of Nash Equilibrium to mixed strategies as follows: the mixed strategy p^* is a Nash Equilibrium if, and only if, for each player i,

$$\mathrm{EU}_i(p_i^*, p_{-i}^*) \geq \mathrm{EU}_i(p_i, p_{-i}^*) \text{ for all mixed strategies } p_i \text{ of player } i.$$

We can identically characterize player i's von Neumann Morgenstern payoff function as

$$\mathrm{EU}_i(p) = \sum p_i(x)\ \mathrm{E}_i(x, p_{-i})$$

where i's pure strategy replaces the mixture.

Thus every mixed strategy Nash Equilibrium has the property that each player's expected equilibrium payoff is the player's expected payoff to any of the actions used with positive probability. For our example of Rock–Scissors–Paper, given the Nash Equilibrium strategy of playing each strategy with 1/3 probability, $\mathrm{EP}(x,\ q^*) = 0$ for x = Rock, Scissors, or Paper, and $\mathrm{EQ}\ (p^*, y) = 0$ for y = Rock, Scissors, Paper, verifying that $p^* = q^* = (1/3,\ 1/3,\ 1/3)$ is a Nash Equilibrium. Nash (1950) demonstrated that every strategic game with a finite number of players with von Neumann-Morgenstern preferences, and a finite number of strategies, has a Nash Equilibrium.

If we modify the Rock–Scissors–Paper game by forbidding DM2 to play Paper, then we have the game depicted in Figure 5.3b. The reader may want to verify that the Nash Equilibrium of this game is $p^* = (2/3, 0, 1/3)$ and $q^* = (2/3, 1/3)$.

Games with Incomplete Information

John Harsanyi (1967/68) formalized the idea of a game with incomplete information. Consider the standard trust game shown in Figure 5.4a. An interpretation of this game is that player 1 chooses between splitting 20 equally with player 2, or sending the 20 to player 2, in which case it becomes 40. Player 2 can then either "cooperate" by choosing to give 15 to player 1 and 25 to herself, or "defect" by keeping the entire 40. The subgame perfect Nash Equilibrium is for player 1 to choose R at n_1 and player 2 to choose R at n_2. That is, if an earnings-maximizing player 2 is given the chance to move, he should "defect" by choosing R. Player 1 anticipates this, and thus chooses not to give player 2 a move, and instead divides the 20 equally.

It is possible that humans have evolved a propensity to cooperate in games like this. For example, people may feel guilty when, as player 2, they defect with the move R. Suppose, furthermore, that guilt reduces a player's final payoff by some amount. A low-guilt person $\mathrm{G_L}$ may only experience a small payoff loss (say 5), while a high-guilt person $\mathrm{G_H}$ may experience a much higher payoff loss (say 20). Finally, we will assume that player 2 knows whether he is a low-guilt or high-guilt type person, but player 1 does not know player 2's type. We can depict this game as a Bayesian trust game with incomplete information, as shown in Figure 5.4b.

There are two important additions to the Bayesian trust game. First, there is a starting move by Nature at node n_0 that determines the type of player 2. Instead of providing a label to the branches that Nature can choose, we indicate the probability that Nature will choose the branch. So with probability 1/4 Nature will move left and the trust game will be played with

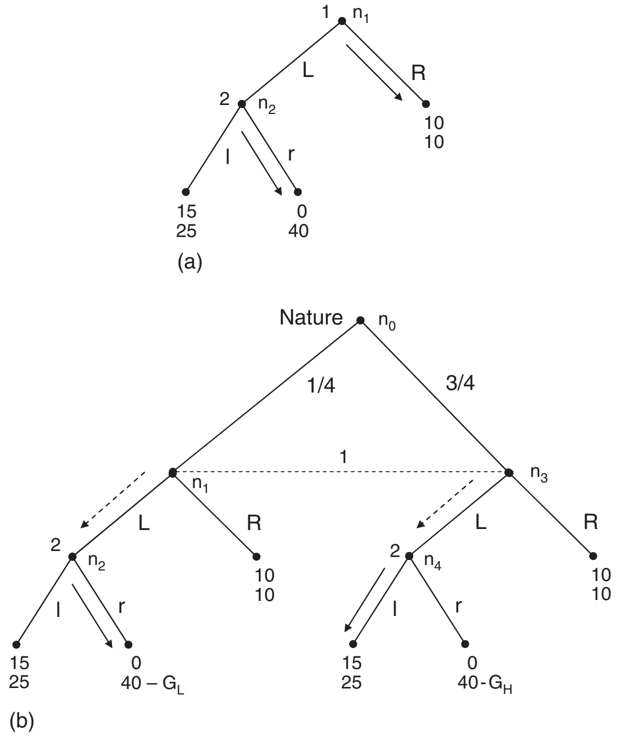


FIGURE 5.4 Bayesian trust game: (a) standard trust game; (b) Bayesian trust game with guilty types, GL = 5, GH = 20. Figure 5.4(a) shows a standard trust game. If DM1 moves left, he is trusting DM2 to move left as well. The only NE of the game is for DM1 to move right and for DM2 to move right. In experiments, this is not what many subjects do. More than 50% of DM1s move left and more than 50% of DM2s reciprocate. One way to explain the experimental results is to use a Bayesian Game, Figure 5.4(b), where some of the DM2s that DM1 encounters feel enough guilt (G_H) to modify DM2s payoff sufficiently to make it rational to reciprocate. Once we allow for types and make the type of DM2 that DM1 faces uncertain, then it is possible to generate predictions similar to the experimental results. Notice this uncertainty is produced by adding a new player called Nature at node n_0, who chooses the type of DM2 (in the case shown, Nature chooses a low-guilt type with probability 1/4 and a high-guilt type with probability 3/4). DM1's uncertainty is shown by the dotted line between n_1 and n_2, indicating that DM1 does not know which of these two nodes he is at when it is time to decide. Notice that while such tricks are easy to produce with prior knowledge of the data, we need better theories of how types emerge and how people form subjective beliefs about types to make such theories predictive.

a low-guilt player 2, and with probability 3/4 Nature will move right and the trust game will be played with a high-guilt player 2. The other important change to the game is the addition of a dotted line indicating that the decision nodes n_1 and n_2 belong to the same *information set*. Up until now players had perfect information about what node they were at, but more general information sets allow us to model a player's lack of knowledge as to where they are in the game. All decision nodes in the same information set must have the same number of branches. In our example, player 1 does not know if he is at n_1 or n_2 when he has to choose L or R, and consequently he cannot make his choice conditional on which node he is at but instead must make the same choice for both nodes.

We can now define a Nash Equilibrium of a Bayesian game as the Nash Equilibrium of the corresponding strategic game where each player is one of the types of one of the players in the Bayesian game. In our Bayesian trust game, that gives us three players (one player 1, one player 2 with low guilt, and one player 2 with high guilt.) Using *Gambit*, we can find four Nash Equilibria for this game, but there is only one subgame perfect Nash Equilibrium, where player 1 always trusts player 2, as shown by the arrows resulting in an expected payoff of 11.25 for player 1. The break-even point for player 1 is a belief that at least 2/3 of the population will feel strong guilt regarding cheating. Thus we can see how optimistic player 1s will try to cooperate by playing, L, while more pessimistic player 1s will play R.

Information sets act as a general formalism for denoting how much information a player has about the previous moves in an extensive form game. Games with perfect information are a special case where all information sets contain only one node, i.e. all players know perfectly what path of the game they are on when it is their turn to make a move. Games with at least one information set containing more than one node are called games with imperfect information. When a player is at an information set, he must act as though he has only one move – that is, the same move at each node in the set – but realize that his move may be along different paths of the game.

A typical game of imperfect information (with many variations) is the simple card game shown in Figure 5.5a. Two players *ante* up by each putting a dollar in the pot. Nature deals one card to player 1. There is a 50–50 chance that the card is high or low. A high card, if shown, wins the pot; a low card, if shown, loses the pot. A fold also loses the pot. At this point, nodes n_1 and n_2, player 1 sees the card and can decide to show (and win a dollar if high, or lose a dollar if low) or raise the bet by adding a dollar to the pot. Now it is player 2's turn to move without knowing player 1's card; thus the information set containing n_3 and n_4. Player 2 can fold (thus losing a dollar) or call by adding a dollar to the pot. A call forces player 1 to show the card, and determines the winner.

A game with imperfect information has perfect recall if every player remembers whatever he knew in the past. We can now distinguish a mixed strategy for an extensive form game as a probability mixture over pure strategies from a behavioral strategy where

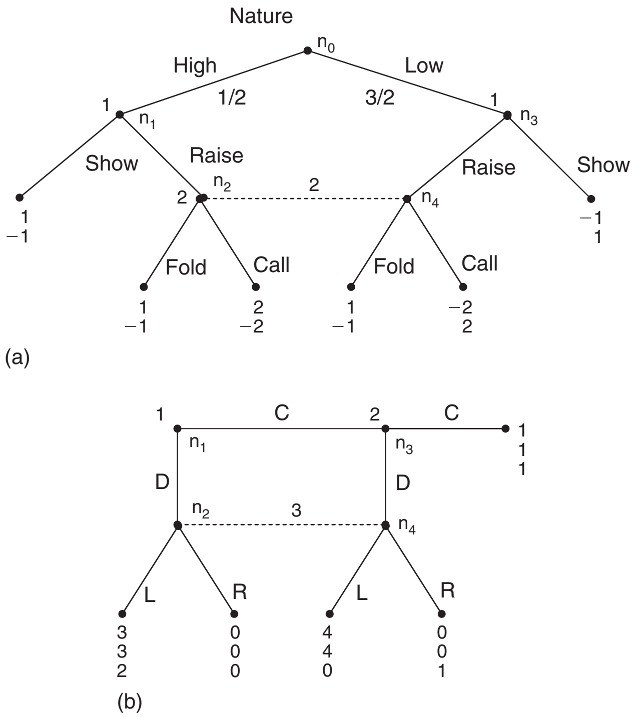


FIGURE 5.5 Imperfect information: (a) simple card game; (b) Selten's Horse. Figures 5.5(a) and 5.5(b) show some simple extensive form games with imperfect information. Again, the imperfection comes from the non-singleton information sets of players denoted by the dotted lines. Figure 5.5(a) shows a simple high–low card game where Nature deals a card to DM1, who sees if the card is high or low and can choose to show his card and win the pot if the card is high, or raise the dollar *ante* by another dollar. DM2 cannot see DM1's card (thus the information set) at n_2 and n_4. Notice the irrelevance of order of the branches at the anode, making it easier to depict the game and decision-makers' information. Figure 5.5(b) depicts Selten's Horse, where DM3 is uncertain if he is at n_2 or n_4. This matters a lot to DM3, who prefers to move L at n_2 and R at n_3. To solve this dilemma, DM3 must take into account the rational behavior of DM1 and DM2. Suppose DM3 believes that DM1 and DM2 may make small mistakes in implementing their strategy. A Nash Equilibrium which is still rational with respect to small mistakes is called trembling hand perfect. In the Horse, (C, C, prop(L) < (1/4)) is trembling hand perfect.

the players pick a probability measure at each information set in the game with the property that each distribution is independent of every other distribution. An important theorem is that any for any mixed strategy of a finite extensive form game with perfect recall, there is an outcome-equivalent behavioral strategy. Note, a behavioral strategy and a mixed strategy are outcome-equivalent if for every collection of pure strategies of the other players the two strategies induce the same outcome. An immediate consequence is the equivalence of Nash Equilibrium for games with perfect recall.

In the card game, player 1's behavioral strategy is two independent probability mixtures over the set {Show, Raise}, while player 2's behavioral strategy is one probability distribution over the set {(Fold, Fold), (Fold, Call), (Call, Fold), (Call, Call)}. The resulting Nash Equilibrium of the game is for player 1 to use the behavioral strategy as follows:

$$\mathrm{prob}(\mathrm{Raise}\,|\,\mathrm{High}) = 1,\ \mathrm{prob}(\mathrm{Raise}\,|\,\mathrm{Low}) = 1/3,$$

$$\mathrm{prob}(\mathrm{Show}\,|\,\mathrm{High}) = 0;\ \text{and},$$

$$\mathrm{prob}(\mathrm{Show}\,|\,\mathrm{Low}) = 2/3.$$

The Nash Equilibrium behavioral strategy for player 2 is prob(Fold) = 1/3, prob(Call) = 2/3. The resulting expected payoff to player 1 is 1/3 while the resulting payoff to player 2 is −1/3.

Trembling Hand Equilibrium

A final game to consider is Selten's Horse, shown in Figure 5.5b. Notice that the pure strategy triple (D, c, L) is a Nash Equilibrium; it does not take into account the possibility that if player 2 actually had a chance to move that player 2 would prefer to move down. Taking this into account, player 3 must play R with at least a 3/4 probability, resulting in another Nash Equilibrium (C, c, $p(\mathrm{L}) \leq 1/4$) and an expected payoff of (1, 1, 1). This equilibrium, which Selten calls a *trembling hand perfect equilibrium*, requires the player's strategy to be optimal even when the other players make small mistakes in their strategies. Formally, a mixed strategy n-tupple is trembling hand perfect if there is a sequence of mixed strategies that converges to the equilibrium strategy and each player's Nash Equilibrium strategy is the best response to all other players' strategies chosen from any n-tupple in the sequence – for example, pick the sequence ($p_\varepsilon(\mathrm{D}) = \varepsilon$, $p_\varepsilon(d) = 2\varepsilon/(1-\varepsilon)$, $p_\varepsilon(\mathrm{R}) = 4/5 - \varepsilon$) and let ε go to 0.

Quantal Response Equilibrium

A final equilibrium concept, due to McKelvey and Palfrey (1995, 1998), is the quantal response equilibrium. A quantal response equilibrium can be defined as follows: let p be a mixed strategy n-vector and let $u'(p) = (u_1'(p), \ldots, u_n'(p))$ be defined by $u_{ij}'(p) = u_i(s_{ij}, p_{-i})$. Notice $u_i(s_{ij}, p_{-i})$ is player i's von Neumann-Morgenstern payoff from choosing the pure strategy s_{ij} given the other players' mixed strategies p_{-i}. Next, define

$$u_{ij}''(p) = u_{ij}'(p) + \varepsilon_{ij}$$

where $\varepsilon_i = (\varepsilon_{i1}, \dots \varepsilon_{im})$ is distributed according to the distribution function $f_i(\varepsilon_i)$ such that the expected value of ε_i is 0. Given $u'(p)$ and f, player i's best response is to choose s_{ij} such that $u''_{ij}(p) \geq u''_{ik}(p)$ for $k = 1, \dots, m$. Note that this relaxes the assumption that players make decisions without errors. In a quantal response equilibrium, best response functions take into account that the game's players make decisions with error: there is some chance that they won't implement a strategy successfully.

Now, given each player's best response function and distribution f_i, calculate $\sigma''_{ij}(u'_i(p))$ as the probability that player i will select strategy j given $u'(p)$. Given f, a quantal response equilibrium for a finite strategic game is a mixed strategy p^* such that for all players i and strategies j, $p^*_{ij} = \sigma''_{ij}(\, u'_i(p^*))$. A quantal response equilibrium exists for any such game. A standard response function is given by the logistic response function

$$\sigma''_{ij}(x_{ij}) = (\exp(\lambda x_{ij})/\textstyle\sum_k \exp((\lambda x_{ik}))$$

$$\sigma''_{ij}(x_{ij}) \text{ where } x_{ij} = u'_{ij}(p).$$

Using *Gambit* and the above logistic response function, we calculate the quantal response equilibrium for Selten's Horse (Figure 5.5b) and see that as $\lambda \to \infty$, $p \to (\mathrm{C}, c, \mathrm{R})$; however, McKelvey and Palfrey (1995) provide a counter-example to show that the limiting quantal response equilibrium does not always correspond to Selten's trembling hand perfection.

It is worthwhile to note that Quantal response equilibrium can be interpreted as a reinforcement learning model (Goeree and Holt, 1999).

GAME THEORY EXPERIMENTS

Game theory experiments test game theory. To be interesting to neuroeconomists typically requires that the experiment goes further than this – for example, by informing the role of emotion in decision or the number of cognitive "types" that exist in a population. This might involve brain-imaging, but need not necessarily do so. In this section we describe the design, practice, and results of game theory experiments that do not include an imaging component but that are especially relevant to neuroeconomic research.

Design and Practice

An important feature of laboratory game theory experiments is that a participant's decisions can be highly sensitive to the specifics of the implementation. An implication is that many game theory experiments are powerful tools for uncovering critical features of the human decision process that might be relatively difficult to detect outside of controlled environments. In addition, like the best theory, the results of game theory experiments might be useful in shedding light on behavioral principles applicable to a large number of decision environments.

The particulars of any experiment design depend on the hypotheses it intends to test. However, there are general design considerations common to any game theory experiment that will be reviewed briefly here. Readers interested in more thorough recent discussions of design and analysis considerations should consult Houser (2008), Camerer (2003, especially Appendix A1.2), Kagel and Roth (1995), Friedman and Sunder (1994), and Davis and Holt (1993). In addition, the outstanding book by Fouraker and Siegel (1963) offers an early but still relevant discussion and defense of procedures in experimental economics. Their work draws attention to instructions, randomization, anonymity, and salient rewards, each of which remains fundamental to experimental economics procedures, as discussed below.

Instructions

It might be expected that decisions would be rather insensitive to the nature of a game's description, so long as the instructions were clear and complete. The reason this is not the case is that instructions not only describe but also frame an environment, and behavior is highly sensitive to framing. For example, using the word "partner" instead of "counterpart" to describe a matched participant in an experiment can affect decisions substantially (see Houser *et al.*, 2004, for further discussion and application to the challenge this raises for interpretation of cross-cultural experiments). As a result, it is important to make an experiment's instructions consistent among various sessions of the same treatment. This is often facilitated by providing a written form of the instructions to subjects, and then reading it to them at the beginning of each session.

Randomization

The role of randomization also cannot be overstated. One reason is that it is necessary for the validity of a variety of widely-used analysis procedures – see, for example, Houser (2008) for elaboration. More generally, the appropriate use of randomization avoids confounding influences on the results. For example, as noted by Fouraker and Siegel (1963), subjects might differ in personality traits or preferences for money, and these differences might also be correlated with the time at which a subject arrives at the experiment. Random assignment of subjects to treatments and roles within the experiment helps to ensure that such differences do not systematically affect an experiment's outcome.

Anonymity

To guarantee anonymity, participants are randomly assigned counterparts, visually separated from each other, and asked to remain silent for the duration of the experiment. By ensuring that participants do not know with whom they are matched, the possibility that decisions will be based on perceptions unrelated to the decision environment under study is largely eliminated. Random and anonymous matching also substantially mitigates the possibility of (unwanted) collusion among participants, which might otherwise confound inferences and interpretations.

It should be emphasized that randomization and anonymity effectively control for differences in individual characteristics only to the extent that each treatment in the experiment uses different participants drawn from a common "cohort," or group with the same distribution of characteristics (e.g., demographic and personality). Often, this is ensured by using a university's general student population to select participants. An alternative is to study the decisions of the same people in multiple treatments. While this can in some cases be efficient, it also raises special design and analysis considerations (e.g., controlling for effects due to the order in which treatments are experienced, as well as the fact that repeated and perhaps correlated observations are obtained from each person).

Salient Rewards

A hallmark of experimental economics, "salient rewards" refer to monetary payments that vary according to a person's decisions in an experiment. Vernon Smith formalized the importance of this procedure with the publication of his Induced Value Theory (Smith, 1976). As long as it is assumed that people prefer more money to less, applying the theory to experiments requires only that real money values are assigned to tokens earned in an experiment. Intuitively, the advantage to doing this is that it raises confidence that participants will recognize the economic incentives implied by the game environment. Induced value theory rigorously justifies experimental tests of game theory, and as such has facilitated the development of new theories incorporating "social preferences" that have been especially relevant to the development of neuroeconomics.

The use of salient rewards in economics experiments stands in sharp contrast to the use of hypothetical rewards common in the psychology literature. As a practical matter, the importance of salient rewards is an empirical question whose answer is likely sensitive to the environment under study (for example, Holt and Laury (2002) compare risk elicitation under both hypothetical and salient reward circumstances). An accepted principle dating from Smith (1965) is that, in relation to hypothetical environments, salient rewards are likely to reduce the variability of decisions among subjects.

Salient rewards are perhaps most transparent in so-called "one-shot" games in which players make a decision, receive their earnings, and then the experiment ends (for example, standard implementations of dictator and ultimatum games). Interestingly, in imaging studies it is typically necessary (for technical reasons) to modify the experiment's protocol so that these games become "multi-shot" or "repeat-single." This means that the game is played several times instead of once, with earnings usually determined by a random draw from one of the completed games. Participants are usually anonymously matched with different players for each game (so-called "Strangers" matching) in order to avoid effects caused by, for example, "reputation," meaning beliefs developed about their counterpart's likely decisions based on play during the course of the experiment.

Experiments with Normal Form Games

Prisoner's Dilemma and Public Goods Games

Prisoner's dilemma (PD) and public goods (PG) games are used to study "social dilemmas" that arise when the welfare of a group conflicts with the narrow self-interest of each individual group member. For example, in a typical two-player PD, each player can choose either to "cooperate" or "defect." Payoffs are symmetric, and chosen so that the sum of the payoffs is greatest when both choose "cooperate" and least when both players choose "defect." However, each player earns the most if he chooses to "defect" when the other cooperates. Thus, the unique subgame perfect Nash Equilibrium of this environment is for both players to defect.

The structure of PG games is similar, but they are typically played in larger groups. In a typical PG game, each member of a group of four people is allocated \$10. Group members simultaneously decide how to allocate their endowment between two "accounts," one private and one public. The private account returns $1 to the subject for each dollar allocated to that account. In contrast, every dollar invested in the public account doubles, but is then split equally among the four group members ($0.50 each). Thus, like the PD game, group earnings are maximized at $80 if everybody "cooperates" and contributes everything to the public account, in which case each of the four participants will earn $20. However, if three subjects contribute $10 each while the fourth "free-rides" and contributes nothing, then the "free-rider" will earn $25 (the best possible outcome for him). Like the PD game, each group member has the private incentive to contribute nothing, and the unique subgame perfect Nash Equilibrium occurs when each subject contributes zero to the group account.

Standard results for PD and PG games are discussed at length elsewhere (see, for example, Davis and Holt, 1993; Ledyard, 1995). The key early finding was that, in aggregate, cooperation occurs about half of the time in PD games, and that about half of the aggregate endowment is contributed to the "public" account in a PG game. It is also routinely found that aggregate cooperation decays when these games are repeated, though cooperation usually remains above zero even with a relatively large number of repetitions (say 30). Though the specific patterns of cooperation can vary with the particulars of the game, the substantive finding that people cooperate in social dilemmas is robust. Results from these early games opened the door for "psychological" game theory (Geanakoplos *et al*., 1989) in which concepts such as reciprocity and altruism play important roles.

PG games continue to be widely studied, and have proven a sharp guide for theories of social preferences (see Chapter 15 of this volume). One reason is that it is simple to develop designs for these games that offer compelling evidence on critical issues in social preference theory. For example, Gunnthorsdottir *et al*. (2007) provide rigorous data from a PG experiment that show that (positive) reciprocity is more important than altruism in driving cooperation. Another reason is that PG games provide rich data on individual decision patterns. For example, PG data starkly reveal that individuals fall into cleanly described "types" (Kurzban and Houser, 2005), and stress that any theory of social preferences that does not account for individual differences is substantively incomplete.

Coordination Games

Unlike standard PD or PG games, many games have multiple equilibria that require coordination. For example, a simple two-player, two-alternative (say "A" and "B") "matching" game will pay each player $1 if they both choose "A" or both choose "B," but will pay each of them nothing if their choices do not match. In these environments, a key role for experiments is to help discover the relative likelihood that a particular equilibrium might be played, as well as the features of the environment (including participant characteristics) that determine this likelihood.

The large literature in coordination games cannot be discussed here, but is well reviewed by Camerer (2003: Chapter 7); this author also suggests several "stylized facts" regarding play in these games. These include that (i) coordination failure is common; (ii) repeated play does not reliably converge to a Pareto-efficient outcome (meaning that no reallocation can make all players simultaneously better off); (iii) the nature of convergence depends on the information available to players and how the players are matched; and (iv) whether and how players are allowed to communicate can have substantial effects on outcomes. Although important challenges arise in its analysis (Houser and Xiao, 2008), communication provides perhaps the richest data for understanding decisions in social environments that require coordination.

Experiments with Extensive Form Games

Ultimatum Games

The ultimatum game, introduced by Guth *et al*. (1982), is a simple take-it-or-leave-it bargaining environment. In ultimatum experiments, two people are randomly and anonymously matched, one as proposer and one as responder, and told they will play a game exactly one time. The proposer is endowed with an amount of money, and suggests a division of that amount between himself and his responder. The responder observes the suggestion and then decides whether to accept or reject. If the division is accepted, then both earn the amount implied by the proposer's suggestion. If rejected, then both the proposer and responder earn nothing.

The key result of ultimatum experiments is that most proposers offer between 40% and 50% of the endowed amount, and that this split is almost always accepted by responders. When the proposal falls to 20% of the endowment it is rejected about half of the time, and rejection rates increase as the proposal falls to 10% and lower. As discussed by Camerer (2003: Chapter 2), ultimatum game results are highly robust to a variety of natural design manipulations (e.g., repetition, stake size, degree of anonymity, and a variety of demographic variables).

An important exception to robustness is reported by Hoffman and Spitzer (1985), who show that offers become significantly smaller, and rejections significantly less frequent, when participants compete for and earn the right to propose. An explanation is that this procedure changes the perception of "fair," and draws attention to the importance of context in personal (as compared to market) exchange environments. These effects might also stem from varying the degree of anonymity among the subjects, or between the subjects and the experimenter (Hoffman *et al*., 1996).

A key focus of recent ultimatum game research has been to understand why responders reject low offers. Economic theory based on self-interested preferences suggests that responders should accept any positive offer and, consequently, proposers should offer the smallest possible positive amount. We review some well-known research on the topic of responder rejections in the "Neuroeconomics experiments" section below.

Trust Games

Joyce Berg, John Dickhaut and Kevin McCabe introduced the popular trust game in 1995. Two participants are randomly and anonymously matched, one as investor and one as trustee, and play a one-shot game. Both participants are endowed with \$10. The investor can send some, all, or none of his \$10 to the trustee. Every dollar sent by the investor is tripled. The trustee observes the (tripled) amount sent, and can send some, all, or none of the tripled amount back to the investor. The amount sent by the investor is a measure of trust; the amount returned by the trustee is a measure of trustworthiness.

Berg *et al*. (1995) reported that investors send about 50% of the endowment on average, and trustees generally return the amount sent. There is more variance in amounts returned than in amounts sent. Indeed, Berg *et al*. (1995) reported that fully 50% of trustees returned \$1 or less. Camerer (1993: Chapter 2) described a variety of studies that replicate and extend these first results. As we discuss further below, this game is also widely used in neuroeconomics experiments.

NEUROECONOMICS EXPERIMENTS

Design

Neuroeconomics experiments provide evidence regarding the biological basis of human decision making. There are many types of neuroeconomic experiments, including (i) purely "behavioral" experiments with healthy volunteers that provide evidence on the role of, for example, emotion on decision; (ii) "lesion" studies that examine the behavioral consequences of brain damage (or temporary disruption with transcranial magnetic stimulation (TMS)); (iii) examinations of drug effects on economic decisions; (iv) skull-surface based measurement of brain electrical activity during decision tasks using electroencephalography (EEG) or magnetoencephalography (MEG); and (v) real-time whole brain imaging using functional magnetic resonance imaging (fMRI) during an economic decision task. A comprehensive review of the leading procedures to draw inferences from brain data can be found in Toga and Mazziotta (2002).

Although each method has unique advantages, over the past decade fMRI has emerged as the dominant technique. The reason is that it is a relatively easily implemented, non-invasive procedure for scientific inference with respect to real-time brain function in healthy volunteers during decision tasks. It is therefore worthwhile to comment briefly on the design and practice of fMRI experiments. Much more detailed discussion can be found in any of a number of recent textbooks that focus exclusively on this topic (Huettel *et al*., 2004 is an excellent and especially accessible source).

Overview

An fMRI neuroeconomics experiment correlates brain activity with economic decision making. However, it does not directly measure neural activity. Rather, evidence on cerebral blood flow is obtained, which Roy and Sherrington (1890) discovered is correlated with underlying neuronal activations. The reason is that active neurons consume oxygen in the blood, leading the surrounding capillary bed to dilate and (with some delay) to an increase in the level of oxygenated blood in the area of neural activity. It turns out that this "hemodynamic response" can be detected and traced over time and (brain) space. Although fMRI technology is rapidly improving, most early studies reported data with temporal resolution of 1 or 2 seconds, with each signal representing a three-dimensional rectangular "voxel" measuring 2 or 3 millimeters on each side and containing literally millions of neurons.

Design

The design of an fMRI neuroeconomics experiment should ensure that the hemodynamic, or blood-oxygen level dependent (BOLD), response can be detected, as well as reliably traced to neural activity associated with the decision processes of interest. A technical constraint in this regard is that the BOLD signal is quite weak, with typical responses being just a few percentage points from baseline measurements made by a typical scanner. An important implication is that neuroeconomic experiments typically require multiple plays of the same game and an averaging of the signals produced therein. That is, single-shot studies are not possible with current technology, and the design strategy must take this into account. A second implication of the weak signal is that other sources of signal variation, such as motion of the subject in the scanner, must be strictly controlled at data collection, and again accounted for during data "preprocessing."

Analysis

The analysis of fMRI data occurs in two stages. The first stage is "preprocessing," the components of which include (i) image realignment to mitigate variation in the data due to head motion; (ii) image standardization to facilitate comparisons among brains of different participants; and (iii) image smoothing to reduce high-frequency voxel specific noise. How different preprocessing approaches affect second-stage inference is the subject of active research (see, for example, Chen and Houser, 2008).

The second stage involves analyzing the (preprocessed) data and drawing inferences about activation patterns. Regardless of the approach used to do this, it is necessary to confront the issue that imaging data has a massive spatial-panel structure: the data include observations from thousands of spatially and temporally characterized voxels. The analysis strategy should allow for the possibility that proximate voxels might have a correlated signal structure, especially because appropriate inference requires accounting for multiple comparisons (see Tukey, 1991, for an accessible discussion of this issue).

Neuroeconomics Experiments with the Trust Game

The Berg *et al.* (1995) trust game (or close variants) has been conducted thousands of times and has played an important role in shaping economists' view of trust and reciprocity. The trust game has also proved a useful paradigm in neuroeconomics. Indeed, it was used by McCabe and colleagues in their 2001 fMRI study of cooperation, which also turns out to be the first published imaging study of economic exchange.

McCabe *et al.* (2001) reasoned that cooperative economic exchange requires a theory-of-mind (ToM). They thus hypothesized that the medial prefrontal cortex, which had been previously implicated in ToM processing (Baron-Cohen, 1995), would also mediate cooperative economic exchange. To test this hypothesis, they asked subjects in a scanner to play variants of a trust game multiple times either with human counterparts outside the scanner or with a computer counterpart. All trust games were "binary" (of the form described by Figure 5.4a), in the sense that both the investor and trustee chose from one of two alternatives, either "cooperate" or "defect." The computer played a known stochastic strategy, and scanner participants were informed prior to each game whether their counterpart was a human or a computer.

Of the study's twelve subjects, seven were found to be consistently cooperative. Among this group, medial prefrontal regions were found to be more active when subjects were playing a human than when they were playing a computer. On the other hand, within the group of five non-cooperators there were no significant differences in prefrontal activation between the human and computer conditions. It is interesting to note that ToM imaging studies caught on quickly, and that the areas identified by McCabe *et al.* (2001) have also been found by others (see Chapter 17 of this volume for a review of the ToM neuroeconomics literature).

Another important imaging (positron emission tomography) experiment with a trust game was reported by de Quervain and colleagues (2004; see also Chapter 15 of this volume). This study sought to provide evidence on the neural basis of punishment, and in particular to investigate whether brain areas including the dorsal striatum are activated when punishing another who has abused one's trust. To assess this, the authors had two anonymous human players, A and B, make decisions in a binary trust game. Both players started with 10 MUs (monetary units), and player A could either trust by sending all 10 MUs to B, or send nothing. If A chose to trust, then the 10 MUs were quadrupled to 40, so that B had 50 MUs and A had zero MUs. B could then send 25 MUs to A, or send nothing and keep the entire 50 MUs. Finally, following B's decision, A could choose to punish B by assigning up to 20 punishment points. In the baseline treatment, each point assigned reduced A's earnings by 1 MU and B's earnings by 2 MUs.

This game was played in a variety of conditions, in order to ensure that the appropriate contrasts were available to assess punishment effects. In addition to the baseline, key treatment variations included the following: (i) a random device determined B's back-transfer, and punishment worked as in the baseline; (ii) B determined the back-transfer, but punishment points were free for A and removed 2 MUs from B's earnings; (iii) B determined the back-transfer, and punishment points were only symbolic in the sense that they were free for A and they also did not have earnings implications for B. With these contrasts in hand, the authors were able to draw the inference that effective (but not symbolic) punishment is associated with reward, in the sense that it activates the dorsal striatum. Moreover, they found that subjects with stronger activations in that area were more likely to incur greater costs in order to punish.

Recently, Krueger *et al.* (2007) have found evidence for two different mechanisms for trust in repeated, alternating-role trust games with the same partner. One system for trust uses the anterior paracingulate cortex in early trials, which is extinguished in later trials and replaced by activation in the septal region of the brain. Bold activations in these areas are interpreted as characterizing a system of unconditional trust in the person. Another system shows no early activation in the anterior paracingulate cortex but does show a late activation consistent with the behavioral responses of subjects to be less trustworthy when temptation is greatest. This is interpreted as characterizing a system of conditional trust, as first movers learn to avoid trusting their partner when temptations to defect are high.

A large number of other trust games have been studied with various motivations. In this volume, trust games play a role in the designs discussed in Chapters 6, 13, 15, 17, 19 & 20.

Neuroeconomics Experiments with the Ultimatum Game

Neuroeconomics experiments with the ultimatum game have been conducted with the primary goal of shedding light on reasons for rejections of unfair offers. Because a person earns a positive amount by accepting the offer, and earns nothing by rejecting, the decision to reject offers has puzzled economists. We here review three innovative studies on this topic, each of which uses a different method: a behavioral study by Xiao and Houser (2005), an fMRI study by Sanfey *et al.* (2003), and rTMS results reported by Knoch *et al.* (2006).

Xiao and Houser (2005) studied the role of emotion expression in costly punishment decisions. A substantial literature suggests humans prefer to express emotions when they are aroused (see, for example, Marshall, 1972). The results obtained by Xiao and Houser (2005) suggest that the desire to express negative emotions can itself be an important motivation underlying costly punishment.

In ultimatum game experiments conducted by Xiao and Houser (2005), responders had an opportunity to write a message to their proposer simultaneously with their decision to accept or reject the proposer's offer. Xiao and Houser found that, compared with standard ultimatum games where the only action responders can take is to accept or reject, responders are significantly less likely to reject the unfair offer when they can write a message to the proposers. In particular, proposers' offers of $4 (20% of the total surplus) or less are rejected 60% of the time in standard ultimatum games. When responders can express emotions, only 32% reject unfair offers, and this difference is statistically significant.

The messages written in Xiao and Houser's (2005) emotion expression game were evaluated using a message classification game with performance-based rewards (for discussion, see Houser and Xiao, 2007, and also the related "ESP" game of von Ahn, 2005). Evaluators were kept blind to the research hypotheses as well as decisions made by participants in the emotion expression game. The vast majority of those who accepted offers of 20% or less wrote messages, and all but one of those messages were classified as expressing negative emotions. An interpretation is that costly punishment decisions occur in part as a way to express dissatisfaction. Earnings maximizing decision making, therefore, is promoted when less expensive channels are available for the purpose of emotion expression.

Sanfey *et al.* (2003; see also Chapter 6 of this volume) is an early fMRI study of the ultimatum game. In this study, participant responders faced either confederate proposers or computers, so that each responder faced exactly the same set of fair (equal split) and unfair offers (between 70% and 90% to the proposer). The brain images revealed that, in comparison to fair offers from human or any computer offers, when the responders were faced with unfair offers from humans there was greater activation in the bilateral anterior insula, the anterior cingulate cortex (ACC), and the dorsolateral prefrontal cortex (DLPFC). The computer condition provides the contrast necessary to rule out the possibility that the source of the activation is the amount of money, thus providing evidence that activations are due to the "unfair" intentions of humans. Moreover, Sanfey *et al.* found that activation in the insula correlated positively with the propensity to reject unfair offers. Because the insula has been implicated in the processing of unpleasant emotions (Calder *et al.*, 2001), this result is convergent evidence that negative emotions underlie the rejection decision in the ultimatum game.

The complexities of the neural networks underlying rejection decisions are underscored by results reported by Knoch *et al*. (2006). These researchers used repetitive transcranial magnetic stimulation (rTMS) in order to disrupt the left or right DLPFC. They found that the rate of rejection of maximally unfair offers (20% was the least amount that could be offered) was just 10% when the right DLPFC was disrupted. On the other hand, the rejection rate of unfair offers was equal to the baseline, 50%, when the disruption was to the left DLPFC. The authors concluded that the right, but not left, DLPFC plays an important role in overriding self-interested impulses, which adds another piece to the puzzle that is the neural underpinning of costly punishment decisions. Other ultimatum game studies are reviewed in various chapters in this volume.

TOWARDS A NEUROECONOMIC THEORY OF BEHAVIOR IN GAMES

Cognitive neuroscience has made great progress regarding the neural basis of perceptual decision making (see, for example, Gold and Shadlen, 2007), as well as value-based decision making (Glimcher *et al*., 2005). Models of decision making based largely on single cell firing in monkeys assumes that neurons encode a sequential probability ratio test (Wald and Wolforwitz, 1947), to decide statistically among competing hypotheses. Within this framework mixed strategies can be explained at the level of neuronal noise (Glimcher *et al*., 2005; Hayden and Platt, 2007), although how noise biases probabilities toward optimal strategies is less understood. It is even less clear how these models of decision making should be extended to games involving other persons.

When individuals evaluate a game tree, they make choices which they expect will result in a desired payoff goal. One approach to solving this problem is to rely on reinforcement learning (Sutton and Barto, 1998) alone, as calculated by the QRE of the game. Such an approach is parsimonious, and would involve only the goal-directed learning parts of the brain (that is, the ventral and dorsal striatum) together with a method for encoding strategies (most likely in the prefrontal cortex) and their payoff equivalents (for example, in pre-motor regions of the brain and the lateral intraparietal area or other parietal areas encoding expected utility maps) (Montague *et al*., 2006). However, one problem with this approach is the relatively long length of time it would take people to learn the QRE of the game. Thus, necessary additions to a reinforcement learning theory of game-playing would be various mechanisms for sharing mental states that would improve the brain choice of an initial strategy and allow the brain to weigh information appropriately and update goals in order to learn more quickly its best strategic choices (starting points for such models might include, for example, Camerer and Ho, 1999, or Erev and Roth, 1998).

Initial strategies are likely to be chosen based on an examination of payoffs leading to a goal set, where a goal set should be understood as the set of all potentially desired outcomes. One unknown is how large a goal set the brain will try to handle. For example, in the game shown in Figure 5.1a, player 1 will see t_1 with a payoff of 50 and the payoff of 40 at t_2 as his goal set from an initial set of payoffs of {50, 40, 30, 20, 0}. In the game shown in Figure 5.4a, player 1 may choose {15, 10} as his goal set from the set of possible payoffs {15, 10, 0}. How players choose their goal sets and edit them over time then becomes a critical feature of such a model. For example, are people more likely to include high payoff outcomes in their initial goal sets?

Given a goals set, a player must identify the paths that will lead to his desired goals. Since each terminal node is isomorphic to a path in the tree, there is a 1–1 and invertible function f which maps the set of goal sets G into the set of game paths P, and therefore there is a set of decision nodes that are "critical" to a player's goals in that at a critical node paths diverge. For example, in Figures. 5.1a and 5.4a, a critical node for player 1 is n_1. Since it is at critical nodes that players make commitments to a proper subset of their goal sets, we expect the brain to weigh the evidence for each path using some form of forward induction and choose based on the resulting accumulation of support for a given strategy.

The next step is to assess who else owns decision rights along the path towards t_1 and what their incentives might be. So, for example, in Figure 5.1 player 2 controls the node n_2 and might like 60 at t_2 compared to 50 at t_1. If this possibility is likely enough, then player 1 may simply decide to play R and get 40. However, player 1 might also try mentally to simulate player 2's mind to induce how player 2 might react at node n_2. Player 1 might reason that player 2 will see that there is a risk to trying for t_3 since player 1 controls the node n_3. But why would there be any risk? A simple answer is that when player 1 took the risk to try for 50, he also made an emotional commitment to punish player 2 if he tried for 60. Notice the decision to punish requires two things; an assessment of shared attention over the fact that player 1 has taken a risk to achieve 50, and an assessment by player 1 that player 2 can empathize with player 1's emotional commitment to punishment.

As part of the forward induction at critical nodes, players are also likely to evaluate the person they are playing as suggested by the nature of the Bayesian trust game shown in Figure 5.4b. In this case, experiential priors from similar situations may bias the players' beliefs (weighing) of the game they are in. When results are evaluated, they will then be updated based on reinforcement learning systems (as a slow learning process) or through much faster emotional responses, such as those found in insula responses.

CONCLUSION

Neuroeconomics research helps to disentangle the complex interrelationships between the neural mechanisms with which evolution has endowed our brains, the mechanisms that our brains have built into our external institutions, and the joint computations of these mechanisms from which social and economic outcomes emerge. Game theory provides a convenient platform for neuroeconomics studies because it formally connects the strategic decisions of multiple individuals to group-level outcomes through a precisely defined mechanism.

We have seen that game theory can entail substantial abstraction. While the level of abstraction created by game theory has substantial advantages, it can also create uncertainty with respect to the way in which laboratory participants perceive the game environment. This can lead to difficulties in interpreting participants' decisions, especially when those decisions are at odds with our constructed notions of rationality. It might be tempting to attribute the failure of a theory to participants' failures to understand the incentives associated with a game. An alternative explanation is that the decisions of "irrational" participants are fully rational, but from an alternative perspective (e.g. "ecological rationality," as promoted by Smith, 2007). Neuroeconomics can help to distinguish between these explanations, and is certain to play a fundamental role in the process of discovering how people decide.

References

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, reciprocity and social history. *Games Econ. Behav.* 10, 122–142.

Bernoulli, D. and Sommer, L. (1954). Exposition of a new theory on the measurement of risk [Specimen theoriae novae de mensura sortis. *Econometrica* 22, 23–36.

Calder, A.J., Lawrence, A.D., and Young, A.W. (2001). Neuropsychology of fear and loathing. *Nature Rev. Neurosci.* 2, 353–364.

Camerer, C. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Camerer, C. and Ho, T. (1999). Experience-weighted attraction learning in normal-form games. *Econometrica* 67, 827–874.

Chen, T. and Houser, D. (2008). Image preprocessing and signal detection in fMRI analysis: a Monte Carlo investigation. George Mason University, Working Paper.

Davis, D. and Holt, C. (1993). *Experimental Economics*. Princeton, NJ: Princeton University Press.

de Quervain, D., Fishbacher, U., Treyer, V. *et al*. (2004). The neural basis of altruistic punishment. *Science* 305, 1254–1258.

Erev, I. and Roth, A.E. (1998). Predicting how people play games: reinforcement learning in experimental games with unique, mixed-strategy equilibria. *Am. Econ. Rev.* 88, 848–881.

Fouraker, L. and Siegel, S. (1963). *Bargaining Behavior*. New York, NY: McGraw-Hill.

Friedman, D. and Sunder, S. (1994). *Experimental Methods: A Primer for Economists*. Cambridge: Cambridge University Press.

Geanakoplos, J., Pearce, D., and Stacchetti, E. (1989). Psychological games and sequential rationality. *Games Econ. Behav.* 1, 60–79.

Glimcher, P., Dorris, M., and Bayer, H. (2005). Physiological utility theory and the neuroeconomics of trust. *Games Econ. Behav.* 52, 213–256.

Goeree, J. and Holt, C. (1999). Stochastic game theory: for playing games, not just for doing theory. *Proc. Natl Acad. Sci. USA*, 96, 10564–10567.

Gold, J. and Shadlen, M. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 22, 535–574.

Gunnthorsdottir, A., Houser, D., and McCabe, K. (2007). Disposition, history and contributions in a public goods experiment. *J. Econ. Behav. Org.* 62, 304–315.

Guth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Org.* 3, 367–388.

Harsanyi, J.C. (1967/68). Games with incomplete information played by "Bayesian" players, Parts I, II, and III. *Management Sci.* 14, 159–182, 320–334, 486–502.

Hayden, B.Y. and Platt, M.L. (2007). Temporal discounting predicts risk sensitivity in rhesus macaques. *Curr. Biol.* 17, 49–53.

Hoffman, E. and Spitzer, M. (1985). Entitlements, rights, and fairness: an experimental examination of subjects' concepts of distributive justice. *J. Legal Stud.* 14, 259–297.

Hoffman, E., McCabe, K., and Smith, V. (1996). Social distance and other-regarding behavior in dictator games. *Am. Econ. Rev.* 86, 653–660.

Houser, D. (2008). Experiments and econometrics. In: S. Durlauf and L.E. Blume (eds), *The New Palgrave Dictionary of Economics*, 2nd edn. Basingstoke: Macmillan.

Houser, D. and Xiao, E. (2008). Classification of natural language messages using a coordination game. Manuscript, George Mason University, Fairfax, VA.

Houser, D., McCabe, K., and Smith, V. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation: comment. *J. Econ. Behav. Org.* 53, 85–88.

Huettel, S., Song, A., and McCarthy, G. (2004). *Functional Magnetic Resonance Imaging*. Sunderland, MA: Sinauer Associates.

Kagel, J. and Roth, A. (eds) (1995). *The Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press.

Knoch, D., Pascual-Leone, A., Meyer, K. *et al*. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314, 829–832.

Krueger, F., McCabe, K., Moll, G. *et al*. (2007). Neural correlates of trust. *Proc. Natl Acad. Sci. USA* 104, 20084–20089.

Kuhn, H.W. (1950). Extensive games. *Proc. Natl Acad. Sci. USA* 36, 570–576.

Kuhn, H.W. (1953). Extensive games and the problem of information. In: H.W. Kuhn and A.W. Tucker (eds), *Contributions to the Theory of Games*, Vol. II. Princeton, NJ: Princeton University Press, pp. 193–216.

Kurzban, R. and Houser, D. (1995). An experimental investigation of cooperative types in human groups: a complement to evolutionary theory and simulation. *Proc. Natl Acad. Sci. USA* 102, 1803–1807.

Ledyard, J. (1995). Public goods: a survey of experimental research. In: J. Kagel and A. Roth (eds), *The Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press, pp. 111–194.

Marshall, J.R. (1972). The expression of feelings. *Arch. Gen. Psychiatry* 27, 789–790.

McCabe, K., Houser, D., Ryan, L. *et al*. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl Acad. Sci. USA* 98, 11832–11835.

McKelvey, R.D. and Palfrey, T. (1995). Quantal response equilibria for normal form games. *Games Econ. Behav.* 10, 6–38.

McKelvey, R.D. and Palfrey, T. (1998). Quantal response equilibria for extensive form games. *Exp. Econ.* 1, 9–41.

McKelvey, R.D., McLennan, A.M., and Turocy, T.L. (2007). *Gambit: Software Tools for Game Theory*, Version 0.2007.01.30 http://gambit.sourceforge.net.

Montague, P.R., King-Casas, B., and Cohen, J.D. (2006). Imaging valuation models in human choice. *Annu. Rev. Neurosci.* 29, 417–448.

Nash, J.F. Jr. (1950). Equilibrium points in N-person games. *Proc. Natl Acad. Sci. USA* 36, 48–49.

Osborne, M.J. (2004). *An Introduction to Game Theory*. Oxford: Oxford University Press.

Roy, C. and Sherrington, C. (1890). On the regulation of the blood supply of the brain. *J. Physiol.* 11, 85–108.

Sanfey, A., Rilling, J., Aronson, J. *et al*. (2003). The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1758.

Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *Intl J. Game Theory*, 4, 25–55.

Smith, V. (1965). Experimental auction markets and the Walrasian hypothesis. *J. Political Econ.* 73, 387–393.

Smith, V. (1976). Experimental economics: induced value theory. *Am. Econ. Rev. Papers Proc.* 66, 274–279.

Sutton, R.S. and Bartow, A.G. (1988). *Reinforcement Learning*. Cambridge, MA: MIT Press.

Toga, A. and Mazziotta, J. (eds) (2002). *Brain Mapping: The Methods*, 2nd edn. London: Academic Press.

Tukey, J. (1991). The philosophy of multiple comparisons. *Stat. Sci.* 6, 100–116.

von Ahn, L. (2005). Games with Purpose. PhD Dissertation, Carnegie Mellon University, Pittsburgh, PA.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*, 2nd edn. Princeton: Princeton University Press.

Wald, A. and Wolfowitz, J. (1947). Optimum character of the sequential probability ratio test. *Ann. Math. Stat.* 19, 326–339.

Xiao, E. and Houser, D. (2005). Emotion expression in human punishment behavior. *Proc. Natl Acad. Sci. USA*, 102, 7398–7401.

CHAPTER

6

Games in Humans and Non-human Primates: Scanners to Single Units

Alan Sanfey and Michael Dorris

OUTLINE

INTRODUCTION

Traditionally, the majority of experimental studies of decision making have examined choices with clearly defined probabilities and outcomes, in which the decision maker selects between options that have consequences for only themselves. The canonical set of decision tasks involves choices between monetary gambles – for example, participants might be asked whether they prefer a 50% chance of $25, or $10 for sure. Though the outcomes and likelihoods are often complex and uncertain, and sometimes ambiguous, these decisions are typically made in isolation.

However, in our daily life decisions are seldom made in these sterile situations, and indeed many of our everyday decisions and choices are made in the context of a social interaction. We live, work, and play in highly complex social environments, and the decisions we make are often additionally dependent on the concomitant decisions of others – for example, when we are deciding to extend an offer of employment or when we are entering a business negotiation. These decisions have the potential to offer a useful window into more complex forms of decision making; decisions that approximate many of the more interesting choices we make in real-life. These types of situation are, however, relatively understudied in the decision-making literature, and thus neuroeconomics has the potential to make important progress in better understanding this class of choices.

The nature of decision making may change fundamentally when the outcome of a decision is dependent on the decisions of others. For example, the standard expected utility computation that underlies many of the existing theories and models of decision making is complicated by the fact that we must also attempt to infer the values and probabilities of our partner or opponent in attempting to reach the optimal decision.

As part of the neuroeconomic approach, several groups of researchers have begun to investigate the psychological and neural correlates of simple social decisions using tasks derived from a branch of experimental economics that focuses on game theory. These tasks, though simple, require sophisticated reasoning about the motivations of other players in the task. The combination of these tasks and modern neuroscientific methods have the potential to greatly extend our knowledge of the brain mechanisms involved in social decision making, as well as advancing the theoretical models of how we make decisions in a rich social environment.

This chapter examines the use of non-invasive imaging techniques in humans and invasive electrophysiological techniques in monkeys for studying decision-making processes during game-theoretic tasks. At the onset, we wish to stress that these are complementary approaches. Each approach has its particular strengths and weaknesses, and each requires that technological hurdles be surmounted, and tasks be modified so they are compatible with these brain-imaging techniques.

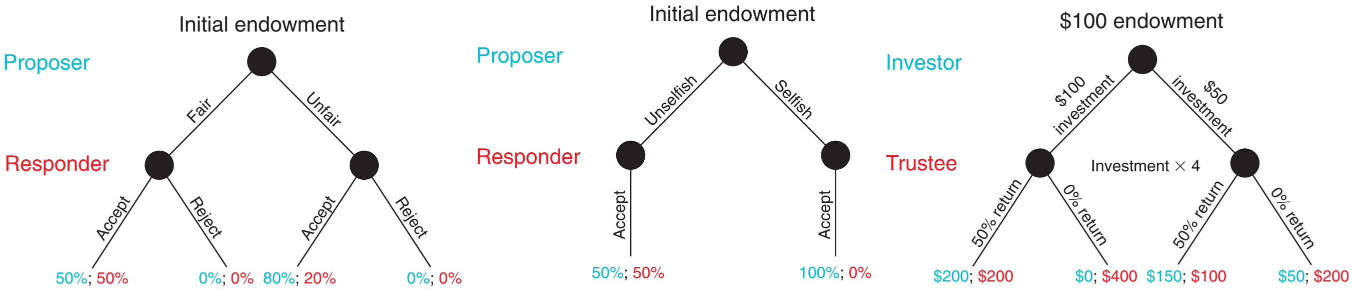
GAME THEORY

In a similar fashion to the framework provided by utility theory for studying individual decisions, game theory offers well-specified models for the investigation of social exchange. The most important development in this field was the work of von Neumann and Morgenstern (1947), whose seminal publication established the foundations of the discipline. In essence, game theory is a collection of rigorous models attempting to understand and explain situations in which decision makers must interact with one another, with these models applicable to such diverse scenarios as bidding in auctions, salary negotiations, and jury decisions, to name but a few.

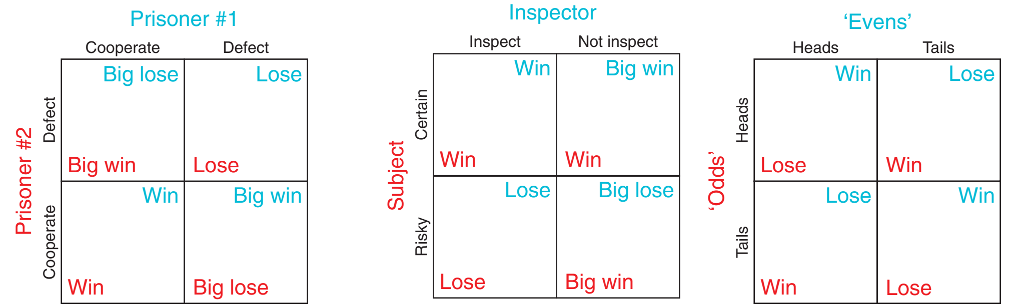
More generally, a common criticism of economic models is that observed decision behavior typically deviates, often quite substantially, from the predictions of the standard model. This is true for the predictions of utility theory for individual decisions (Kahneman *et al.*, 1982), as well as game theory for social decisions. Classical game theory predicts that a group of rational, self-interested players will make decisions to reach outcomes, known as Nash Equilibria (Nash, 1950), from which no player can increase her own payoff unilaterally. However, ample research has shown that players rarely play according to these strategies (Camerer, 2003). In reality, decision makers are typically both less selfish and more willing to consider factors such as reciprocity and equity than the classical model predicts.

Nonetheless, the well-characterized tasks and formal modeling approach offered by game theory provides a useful foundation for the study of decisions in a social context. From an experimental standpoint, the mathematical framework of game theory provides a common language in which findings from different research groups, and indeed research methodologies, can be compared, and deviations from model predictions quantified. These tasks produce a surprisingly varied and rich pattern of decision making, while employing quite simple rules (Figure 6.1 provides a useful summary of standard tasks; see Camerer, 2003 for a summary of results). Importantly, behavioral and neurobiological studies of social decision making are also proving instructive in understanding the nature of the discrepancies between model predictions and observed behavior.

One common focus of game theory is bargaining behavior, with the family of dictator and ultimatum games often used to examine responses to equality and inequality. In the dictator game (DG), one player (the proposer) decides how much of an endowment to award to the second player (the responder). Allocations in this game measure pure altruism, in that the proposer sacrifices personal gain to share some amount of the endowment with the responder. The ultimatum game (UG) (Guth *et al.*, 1982) examines strategic thinking in the context of two-player bargaining. In the UG, the proposer and responder are also asked to divide a sum of money, with the proposer specifying how this sum should be divided between the two. In this case, though, the responder has the option of accepting or rejecting the offer. If the offer is accepted, the sum is divided as proposed. However, if it is rejected, neither player receives anything. In either event the game is over; that is, there are no subsequent rounds in which to reach agreement. If people are motivated purely by self-interest, the responder should accept any offer and, knowing this, the proposer will offer the smallest non-zero amount. However, this Nash Equilibrium prediction is at odds with observed behavior and, at least in most industrialized cultures, low offers of less than 20% of



(a) *Ultimatum Game* (*UG*) (b) *Dictator Game* (*UG*) (c) *Trust Game* (*TG*)



(d) *Prisoner's Dilemma Game* (*PDG*) (e) *Inspection Game* (*IG*) (f) *Matching Pennies Game* (*MPG*)

FIGURE 6.1 Outline of the structure of several standard game theoretic tasks. In the bargaining tasks (a, b, c), the initial endowment provided varies across studies, and the proposer/investor are free to offer any amount of this investment – sample amounts have been shown for illustrative purposes. These games are typically sequential, with the proposer/investor making an offer, and then the responder/trustee responding in turn. For the competitive games (d, e, f), the two players generally make simultaneous decisions, with the monetary payoffs also varying across studies, though they broadly correspond to the outcomes shown.

the total amount are rejected about half of the time. There are some interesting differences in more traditional cultures (Henrich *et al.*, 2005), but in general the probability of rejection increases substantially as offers decrease in magnitude. Thus, people's choices in the UG do not conform to a model in which decisions are driven purely by self-interest, and, as will be discussed below, neuroscience has begun to offer clues as to the mechanisms underlying these decisions.

In addition to bargaining, reciprocal exchange has been studied extensively in the laboratory, exemplified by trust and closely-related prisoner's dilemma games. In the trust game (TG), a player (the investor) must decide how much of an endowment to invest with a partner (the trustee) in the game. Prior to this investment being transferred to the trustee, the experimenter multiplies this money by some factor (usually tripled or quadrupled), and then the trustee has the opportunity to return some or all of this increased amount back to the investor, but, importantly, need not return any money if she decides against it. If the trustee honors trust, and returns money to the investor, both players end up with a higher monetary payoff than was originally obtained. However, if the trustee abuses trust and keeps the entire amount, the investor ends up with a loss. As the investor and trustee interact only once during the game, game theory predicts that a rational and selfish trustee will never honor the trust given by the investor. The investor, realizing this, should never place trust in the first place, and so will invest zero in the transaction. Despite these rather grim theoretical predictions, in most studies of the TG a majority of investors do in fact send some amount of their money to the trustee, with this trust typically reciprocated.

The well-studied prisoner's dilemma game (PDG) is similar to the trust game, except that in the standard version both players now simultaneously choose whether or not to trust each other, without knowledge of their partner's choice. In the PDG, the players each choose to either cooperate or not with their opponent, with their payoff dependent on the interaction of the two choices. The largest payoff to the player occurs when she defects and her partner cooperates, with the worst outcome when the decisions are reversed (player cooperates while partner defects). Mutual cooperation yields a modest payoff to both players, while mutual defection provides a lesser amount to each. The Nash Equilibrium for the PDG is mutual defection, which, interestingly, is in fact a worse outcome for both players than mutual cooperation, but again, in most iterations of the game players exhibit much more trust than expected, with mutual cooperation occurring about 50% of the time.

Public goods games are a generalized form of the PDG, with each player able to invest a proportion of an endowment provided by the experimenter in a public good, which is then increased in value and shared back with all players. The self-interested solution here is to hold back on investment and hope that everyone else contributes the maximum amount, modeling situations such as environmental pollution or upkeep of a public park. However, as in PDG cases, players on average contribute about half of their endowment to the public good.

Finally, games that typically call for mixed strategy equilibrium solutions, such as matching pennies and the inspection game, offer insights into how we assess the preferences of others and choose accordingly. For example, in matching pennies, each player chooses between two alternatives (such as heads or tails). One player (evens) wins if the two choices are the same, and the other (odds) wins if they are not. The Nash Equilibrium is to select the two alternatives randomly with equal probabilities, but players typically approach this game by attempting to infer the strategy of our opponent, thus providing a window into how we use theory-of-mind processes to assist our strategic decision making.

Of course, in many of the cases discussed here, such as fair offers in the UG, cooperation in PDG, and contribution in PG experiments, it is unclear whether the decisions emerge from strategic or altruistic motivations. Do I offer you 50% of the pot in an ultimatum game because I value fairness, or because I fear you will reject anything less? Examining these games in a neural context can begin to offers clues as to the motivations behind the decisions, and the combination of game theory and neuroscience therefore offers a useful set of tasks, a rigorous mathematical modeling approach, and techniques to allow us to begin probing the underlying processes of social decision making.

Recent research has combined these behavioral paradigms from experimental economics with a variety of methods from neuroscience in an effort to gain a more detailed picture of social decision making. The benefits of this approach are twofold. First, as described above, actual decision behavior in these tasks often does not conform precisely to the predictions of classical game theory, and therefore more precise characterizations of behavior, in terms of the neural and psychological process that underlie them, will be important in adapting these models to better fit how decisions are actually made. Secondly, neuroscience can provide important biological constraints on the processes involved, and indeed research is revealing that many of the processes thought to underlie this type of complex decision making may overlap strongly with more fundamental brain processes such as reward, disgust, pain, etc. Knowledge of the "building blocks" of decision making in games will greatly assist in constructing better models of this process.

GAMES IN NON-HUMAN PRIMATES

Although use of awake, behaving monkeys has been a mainstay of systems neuroscience research for over 40 years, their use in conjunction with game-theoretic tasks is less than 5 years old. Though still in its infancy, this research has already produced significant insights into the hidden processes that occur within the so-called "black box" during social interactions. Here we outline the current state of this research, not only to illustrate how specific studies have advanced our understanding, but also to highlight the promise (and limitations) of these neurophysiological techniques in providing future insights.

The Animal Model and Sensory-motor System

A suitable animal model is required to permit direct access to the neural substrate during game play. For a number of reasons, the rhesus monkey (*Macaca mulatta*) has been the primary animal model for studying higher-order decision processes. The general organization of their nervous system is similar to that of humans, with this complexity allowing these non-human primates to learn relatively sophisticated behavioral tasks. Across a number of decision-making contexts, including that of mixed-strategy games on which we focus here, monkeys and humans display comparable strategies, suggesting that many of the underlying neural processes are shared.

For a number of practical reasons, decision-making research has focused primarily, but not exclusively (Kalaska *et al.*, 2003; Romo and Salinas, 2003), on the monkey visuosaccadic system (Schall and Thompson, 1999; Glimcher, 2003). The visuosaccadic system is of critical importance because it allows us to efficiently extract visual information from our environment.

It achieves this by alternating between periods of fixation, when visual information is acquired by the retinas and processed in extra-striate visual areas, and ballistic eye movements known as saccades which align the high acuity foveae on targets of interest. Although not traditionally considered "choices," saccades are in fact the behavioral read-out of one of our most common decisions – that of choosing when and where to look.

The neural circuitry underlying visual processing and saccadic control is well understood, which provides a solid foundation for asking questions about the decision processes that link sensation to action. The simplicity of saccades aids in this understanding; three pairs of antagonistic eye muscles move a relatively inertia-free globe in a stereotyped manner. Attributing neuronal activity to the movement of other motor effectors is complicated by the complex interactions that occur between multiple muscles across many joints, and the variable loads and dynamics associated with these movements. Finally, the visuosaccadic neural circuitry is housed entirely within the cranium, thus providing the stability necessary for recording tiny neurons within an awake and moving preparation.

A critical feature of visuosaccadic neurons that must be understood in order to interpret neurophysiological decision studies is that of the response field. Each visuosaccadic neuron is activated by a particular combination of sensory and motor attributes which together define the neuron's response field. Populations of neurons with similar response fields are organized together into topographic maps of sensory and motor space. Sensory attributes may include the spatial location of visual stimuli relative to the foveae, the speed and direction of motion, and color and shape. Motor attributes may include the direction and amplitude of the saccadic vector and the timing of the response. Therefore, the sensory and motor attributes of each neuron are typically determined at the onset of an experiment so that decision tasks can be tailored to robustly activate the neuron under study.

Response fields are transformed in two ways that are relevant to the decision-making process. First, response-field properties evolve as we move from sensory- to motor-related brain regions; early on, response fields encode sensory properties largely irrespective of motor responses, and later on, response fields encode properties of the movements largely irrespective of incoming sensory attributes. This visuomotor transformation has been well characterized by decades of neuroscience research. Second, response field activation is shaped by cognitive and economic factors even when immediate sensory and motor attributes are fixed. These modulatory processes result from interactions with different regions of the visuosaccadic network and with brain regions that lack classical response fields, such as much of the frontal cortex and basal ganglia. A neuroeconomic approach promises to advance our understanding of how neuronal response fields are transformed by such contextual information.

Advantages and Disadvantages of a Systems Neurophysiology Approach

The advantages of the systems neurophysiology approach stem from the direct access to the neural substrate that it provides. Neuronal signals can be sampled with exquisite temporal (<1 ms) and spatial (individual neurons) resolution and, with nearly comparable precision, neuronal activity can also be artificially manipulated.

For those not familiar with the methodology associated with neurophysiology in awake, behaving monkeys, we will outline it briefly. To gain access to the neural structures of interest, a surgical craniotomy is performed which involves drilling a hole in the skull while the monkey is under general anesthesia. A chamber with a removable cap is fixed over this craniotomy and cleaned daily under antiseptic conditions. At the onset of each experiment, a fine metal electrode or needle pierces the membranes which cover the brain and, with high precision, is slowly lowered to the brain region of interest. These procedures are painless and cause little damage to neural tissue, because the brain lacks pain receptors and only very thin probes are used. These latter properties are critical, because to obtain accurate experimental results both the animal and brain must be in as natural a state as possible.

It is the action potentials, or electrical pulses originating in one neuron and propagating along extended processes to communicate with other neurons, that are recorded with microelectrodes during these monkey experiments (see Figures 6.3 and 6.4, later in this chapter, for examples). Sampling the activity of individual neurons over many experimental sessions provides a statistical approximation of the role of a specific brain region in the decision process. For example, neuronal activity can be correlated to features of the sensory instructions, internal variables predicted by economic theory, aspects of the choice response, and the type of reinforcement. Because this neural activation can be measured with millisecond precision, it is the best means for understanding the moment-to-moment computations that underlie the decision process.

Artificial manipulation of neuronal activity can provide causal evidence that a brain region is involved in the decision process, complementing the correlational evidence provided by neuronal recordings. One way to manipulate neuronal circuits is to inactivate a brain region. Inactivation can be either permanent, through surgical or chemical lesions, or temporary, through the injection of pharmacological agents or physical cooling. Another way to artificially manipulate neuronal activity is through electrical micro-stimulation. Microstimulation excites neuronal tissue, and its temporal precision, spatial extent, and intensity can be controlled more precisely than with inactivation techniques.

A number of potential disadvantages exist in using non-human primates to infer the neural processes underlying human social interactions. To date, non-human primates have only been trained to perform simple mixed-strategy games. Monkeys may not be a suitable animal model for more sophisticated games, such as UG and PDG, because they may lack key cognitive abilities found in humans. Moreover, it may difficult to train animals on game-theoretic tasks without verbal instructions and using only operant conditioning techniques. Even if comparable choice strategies are used during game play, we must remember that this is a prerequisite, not proof, for the same neural mechanisms being shared in these two species. That being said, monkeys and humans have displayed remarkably similar strategies under the simple mixed-strategy games studied to date (Barraclough *et al.*, 2004; Dorris and Glimcher, 2004; Lee *et al.*, 2004, 2005). Although it remains to be seen what the limits of this animal model will be, understanding the neural mechanisms underlying game play in monkeys is important because these may be directly related to our own decision-making mechanisms; at the very least, they represent the core mechanisms upon which our more sophisticated decision processes rest.

Adapting Games for Non-human Primates

Neurophysiologists have initially focused their efforts on simple mixed-strategy games primarily because non-human primates can be trained relatively easily on these tasks. We will briefly describe some of these games and how they have been modified for the neurophysiology laboratory (see Figures 6.3a and 6.4a). The reader is also referred to Chapter 31.

All tasks to date involve thirsty animals competing against dynamic computer opponents for liquid rewards. At the onset of each experiment, a microelectrode is manipulated to isolate the activity of a single neuron from background brain activity. Before game play begins, the experimenter typically determines the neuron's response-field properties, as described previously, and tailors the choice targets so that the neuron under study is maximally activated. Each game trial begins with the animal fixating a central visual stimulus. The animal indicates its choice by directing a saccade to one of the peripheral targets upon their presentation. Whether the animal receives a liquid reward depends on both its own choice and that of the computer opponent. Although computer algorithms vary in their details across studies, all look for patterns in the animal's history of choices and rewards in an effort to predict and counter the animal's upcoming actions.

Monkeys have been trained to perform simple zero-sum games such as "matching pennies" (Barraclough *et al.*, 2004; Lee *et al.*, 2004) and "Rock–Paper–Scissors" (Lee *et al.*, 2005) and non-zero-sum games such as the "inspection game" (Dorris and Glimcher, 2004), in which the Nash Equilibrium solution varies from block to block. Another successful means for studying adaptive decision making in non-human primates uses "matching law tasks," in which the allocation of responses is proportional to the relative reinforcement schedules associated with the available responses (Sugrue *et al.*, 2004; Corrado *et al.*, 2005; Lau and Glimcher, 2005). Because matching law tasks do not involve interaction with a strategic opponent they are technically not games; however, we include them here because it is unclear whether monkeys can distinguish between these two classes of adaptive tasks. Chapter 30 of this volume describes neurophysiological work employing matching tasks.

Below, we describe recent insights provided by neurophysiological approaches in simple mixed-strategy games. These experiments examine activation across a wide range of the visuomotor network, including the parietal cortex, brainstem, and frontal cortex. Broadly speaking, these experiments examine how the desirability of sensory stimuli is encoded, motor actions are selected and the consequences of these actions are evaluated, respectively, during mixed-strategy games (Figure 6.2).

Encoding the Desirability of Choice Stimuli in Parietal Cortex

First, we address how the representations of sensory stimuli are influenced by the subjective desirability of their associated actions. The lateral intraparietal area (area LIP) is a region of the parietal lobe important for the decision-making process because it is situated at the end of visual processing stream, and its outputs impact regions of the brain involved in

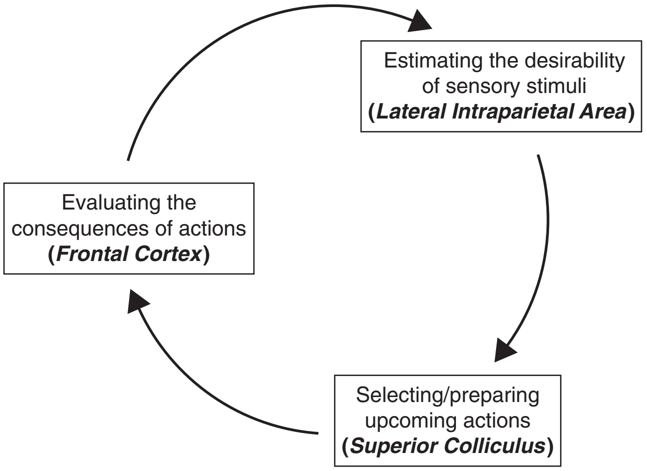


FIGURE 6.2 Schematic of important neural structures studied using mixed strategy games in non-human primates.

planning and executing upcoming saccades (Pare and Wurtz, 2001; Bisley and Goldberg, 2003; Grefkes and Fink, 2005). Previous work demonstrated that activity in this region may encode the saliency of visual targets in a manner that can be used to allocate attentional resources and/or to select between upcoming saccade goals (Andersen, 1995; Goldberg *et al.*, 2006). A pioneering study conducted by Platt and Glimcher (1999) demonstrated that important variables predicted by economic theory, such as the probability and magnitude of reward, impact the firing rates of LIP neurons and, in doing so, provided an alternative decision theory framework for studying the role of brain regions in simple sensory-to-motor transformations.

Given that area LIP lies at the nexus between sensory and motor processing and is influenced by economic variables, Dorris and Glimcher (2004) hypothesized that it could play an important role in representing the desirability of potential choice targets under game conditions. Monkeys competed against a computer opponent during the mixed-strategy inspection game (Figure 6.3a). The payoff matrix was experimentally manipulated across blocks of trials so that the Nash Equilibrium solution for the monkey ranged from choosing the target in the center of the neuron's response field 10% of the time to choosing it 90% of the time. If LIP encoded the probability of movement, its activation would vary across blocks of trials. If, however, LIP encoded the desirability of the target stimulus, its activation should remain relatively constant. This latter interpretation is an extension of the Nash Equilibrium concept which suggests that the subjective desirability is, on average, equal between the available options during mixed-strategy games. LIP activity was indeed shaped by the subjective desirability of choice stimuli; firing rates varied along with changing desirability under forced-choice conditions (Platt and Glimcher, 1999; Dorris and Glimcher, 2004) (Figure 6.3b) and remained constant throughout the behavioral equilibria established during mixed-strategy conditions (Dorris and Glimcher, 2004) (Figure 6.3c).

Although the Nash Equilibrium concept posits that there is no incentive for an individual to change her overall strategy once at behavioral equilibrium (Nash, 1950), it is still possible that internal representations of desirability are biased towards particular options from trial to trial (Harsanyi, 1974). The precise signals obtained from recording single neurons make this an ideal technique for examining any subtle fluctuations in desirability. To estimate desirability on a trial-by-trial basis, Dorris and Glimcher (2004) optimized a simple reinforcement learning algorithm to the monkey's pattern of behavioral choices using maximum likelihood methods. Briefly, the desirability of each target was incremented if reward was received for choosing the risky option, or decremented if reward was withheld for choosing the risky option. The only free parameter was the "learning rate" at which desirability was updated based on this reward information. The iterative nature of this reinforcement learning algorithm resulted in an estimate of desirability derived from all the subject's previous choices, with the most recent choices being weighted most heavily. Indeed, trial-by-trial fluctuations in LIP activity co-varied with this trial-by-trial behavioral estimate of subjective desirability (Dorris and Glimcher, 2004) (Figure 6.3d).

Similar LIP recording experiments have also been conducted using a "matching-law" task (Sugrue *et al.*, 2004). To estimate the subjective desirability of responses on each trial, these experimenters used a function that weighted local reward history in a manner that closely approximated the iterative algorithms associated with reinforcement learning. They also found that LIP activation remained constant on average, and that a local estimate of reward rates was predictive of choices on a trial by trial basis (Sugrue *et al.*, 2004, 2005). At this early stage, the specific form and parameters of modeling efforts will surely be refined with further experimentation (see Chapters 22, 24, 30, 31, and 32 for further advances in modeling techniques). More generally, what these experiments demonstrate is that the high fidelity neuronal signals afforded by recording single neurons allow for moment-to-moment correlations between neuronal activity and behavioral responses, thus providing unprecedented insight into the neuronal mechanisms underlying stochastic choice.

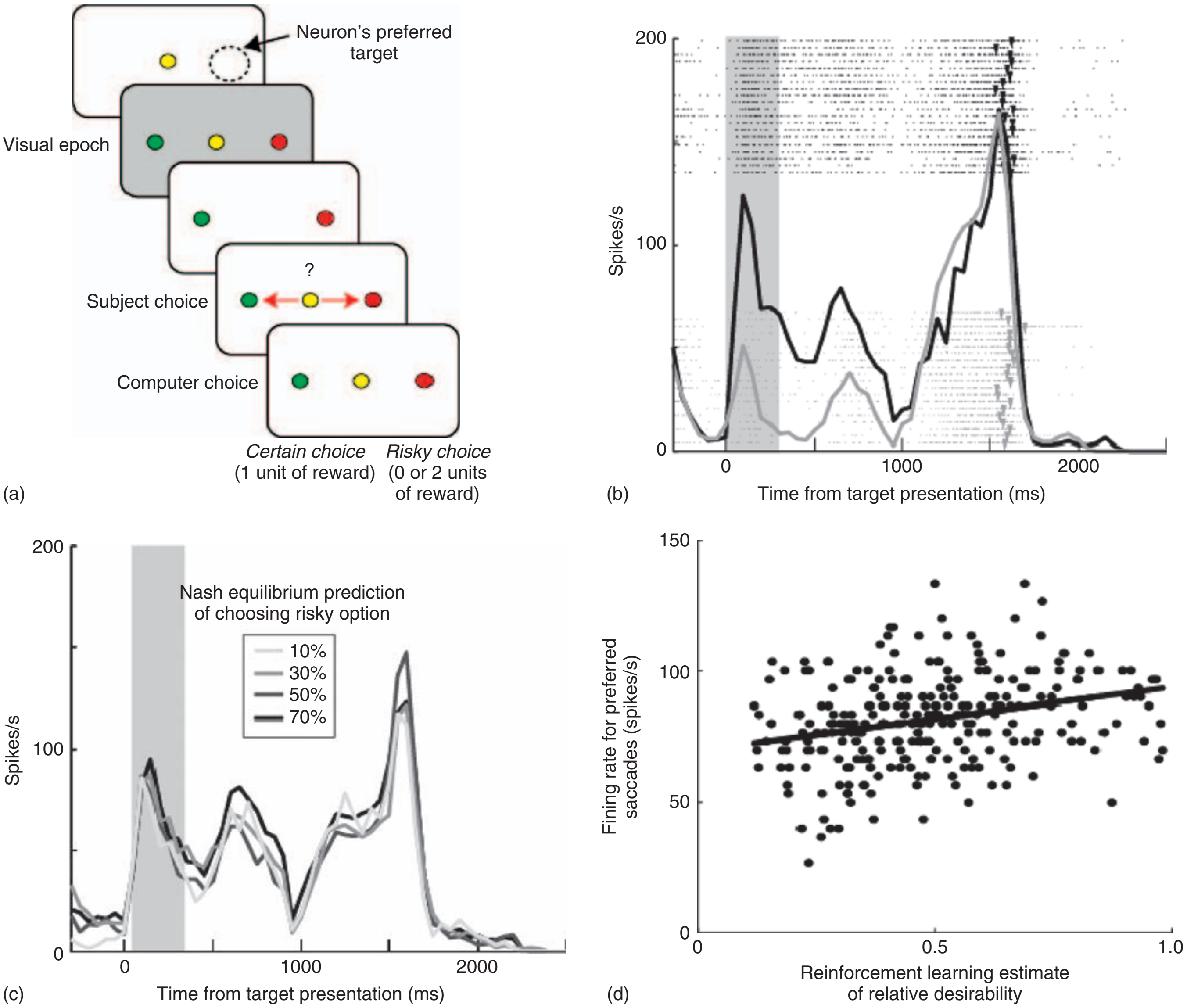


FIGURE 6.3 Encoding the subjective desirability of visual targets in area LIP. (a) Visuosaccadic version of mixed-strategy inspection game. (b) The activity of a single LIP neuron during an instructed task. Initial visual responses where influenced by the desirability of the neuron's preferred target are shaded gray. Black line = two-thirds of total reward associated with preferred target; gray line = one-third of total reward associated with preferred target. (c) Activity of same neuron during mixed-strategy inspection game. Despite changes in the probability of preferred responses, LIP activity remained relatively constant, which is consistent with an overall equivalency in desirability at mixed-strategy equilibria. (d) Trial-by-trial variability in activity during the visual epoch was significantly correlated to a behavioral estimate of desirability. Adapted from Dorris and Glimcher (2004).

Evolving Response Selection in Midbrain Superior Colliculus

Although area LIP appears to represent the desirability of visual stimuli, which is of critical importance for selecting upcoming saccades, it contributes little to the actual generation of saccadic movements themselves. This is evidenced by the large currents required to trigger saccades with micro-stimulation in area LIP, the poor correlations of LIP activity with saccadic reaction times, the relatively mild effects on saccade generation resulting from its ablation, and the simple fact that LIP requires visual inputs for robust activation (Goldberg *et al.*, 2006). The midbrain superior colliculus (SC), by contrast, is intimately involved in saccade generation; saccades are evoked with micro-stimulation at low currents, activity patterns are predictive of both when and where a saccade will occur, and, anatomically, it provides the main drive to the saccade burst generator in the brainstem

(Robinson, 1972; Glimcher and Sparks, 1992; Dorris *et al.*, 1997; Grantyn *et al.*, 2004). This section examines how activity within the SC evolves to select one saccade response over another during the mixed-strategy game "matching pennies" (Figure 6.4a).

On each trial of matching pennies, both the monkey and the computer opponent selected one of the two available targets. The monkey received a liquid reward if it chose the same target as the computer, and nothing otherwise. Monkeys approach the matching pennies Nash Equilibrium solution of choosing each of the two options stochastically and in equal proportions (Lee *et al.*, 2004). Although behavior is relatively unpredictable, examination of SC neuronal activity reveals that one saccade becomes increasingly selected over the other as the time of target presentation approaches (Figure 6.4b). Therefore, the degree to which neuronal activations segregate over time provides insight into the time course of response selection preceding strategic actions.

Direct perturbation of neural circuits has been used in decision tasks to provide functional evidence regarding the contribution of a brain region to choice behavior (Salzman *et al.*, 1990; Gold and Shadlen, 2000; Carello and Krauzlis, 2004; Dorris *et al.*, 2007). Here, a micro-stimulation paradigm adapted from Gold and Shadlen (2000, 2003) tested whether the predictive activity in the SC outlined above is functionally related to the process of response selection under game conditions. On a small proportion of matching pennies trials, the ongoing decision process was perturbed with a short burst of micro-stimulation (Figure 6.4c). This stimulated SC location elicited saccades orthogonal to the direction of the choice targets. Because saccade trajectories are determined by population activity across the topographically organized SC map (Lee *et al.*, 1988), stimulation-induced saccades deviate towards regions of pre-existing activity. Indeed, these stimulation-induced saccades deviated towards the location the animal ultimately chooses (Figure 6.4c). Interrupting developing saccade plans at a range of times preceding the presentation of the choice targets thus opens a window into the time course of the response selection process (Figure 6.4d). These results highlight how artificially perturbing activity within decision circuits can provide insight into the functional role that a particular brain region plays in the decision process.

Evaluating the Consequences of Actions in Frontal Cortex

The final significant work involving game play in non-human primates demonstrates that the frontal cortex contains signals that could be used to evaluate the consequences of actions during game play. Actions and their associated payoffs must be tracked in order for an agent to adapt their choice strategies during social interactions. Previous work has demonstrated that activity throughout the basal ganglia and frontal cortex is sensitive to reinforced actions under pure-strategy conditions as animals learn to converge on a single correct option (Schultz *et al.*, 2000; Balleine *et al.*, 2007). What remained unclear was how action value representations were updated under mixed-strategy conditions when there is no single correct answer and agents must respond stochastically from trial to trial.

Daeyeol Lee's group has demonstrated that the firing rates of individual neurons in the dorsolateral prefrontal cortex (dlPFC) are sensitive to both the particular choice (i.e., left vs right) and the consequences of those choices (i.e., rewarded or unrewarded) during the matching-pennies game (Barraclough *et al.*, 2004). Moreover, certain neurons were preferentially activated by particular combinations of choices and rewards (e.g., left and unrewarded), suggesting that the dlPFC may also be involved in integrating these two sources of information. Activity within another frontal region, the dorsal anterior cingulate cortex (dACC), encoded critical information about the temporal delay of previous rewards within a sequence of responses (Seo and Lee, 2007). Recently, these researchers have begun to use these neural signals as the inputs for reinforcement learning algorithms to predict choice patterns during mixed-strategy games. See Chapter 31 for further details of this modeling work.

This work further demonstrates the promise that direct recording of neural signals has for uncovering the mechanistic algorithms underlying stochastic choice. This work is also noteworthy because it illustrates the importance of recording the spatial resolution of individual neurons. Neural structures close to the sensory input or motor output are generally organized topographically, with large populations of neurons firing together under similar conditions. Most association areas involved in the decision processes, such as the dlPFC and dACC, are not organized in a topographic manner. Instead, neurons performing abstract calculations within the decision process, such as tracking specific combinations of actions and rewards or rewards within a particular sequence, are intermingled throughout these brain areas. Under these circumstances, even those non-invasive imaging techniques with relative high spatial resolution, such as fMRI ($\sim$1 mm^3), may have difficulty detecting these distributed signals.

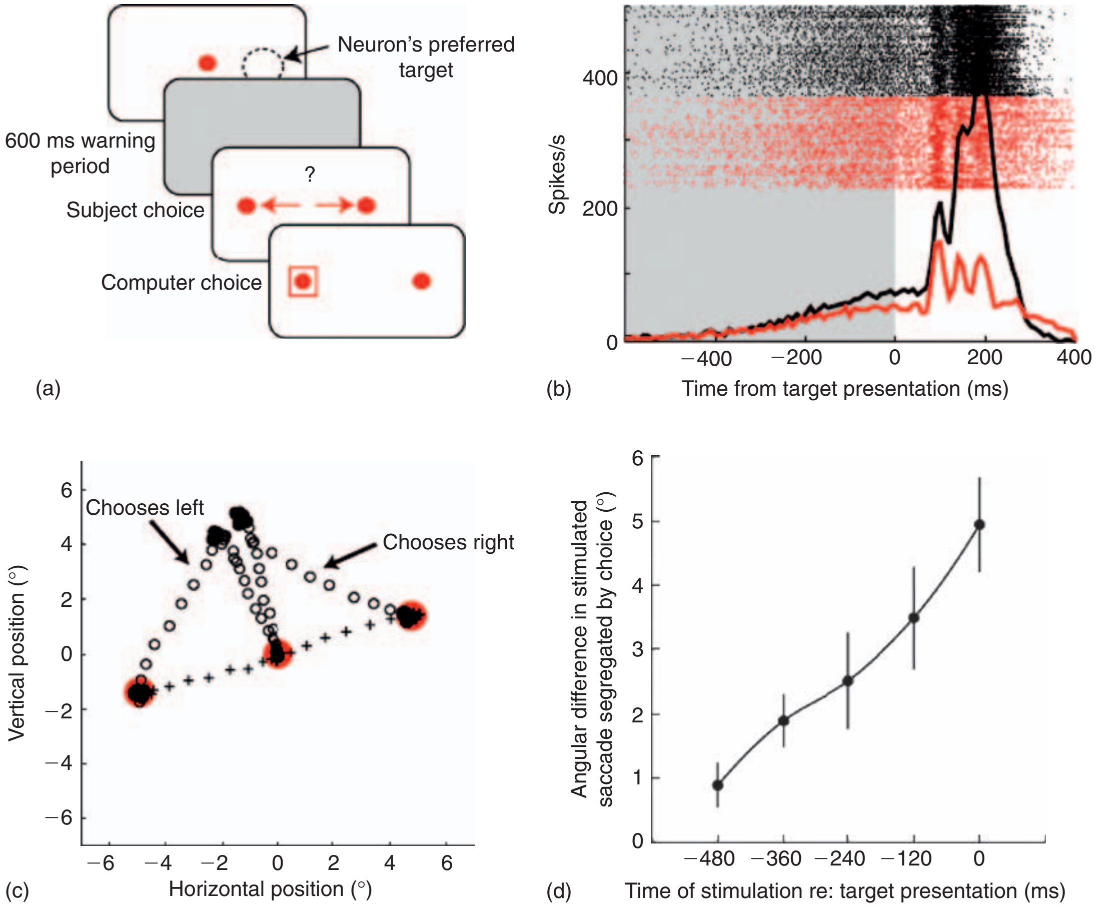


FIGURE 6.4 Preparing upcoming actions in midbrain SC. (a) Visuosaccadic version of mixed-strategy matching-pennies game. Analysis is focused on the warning period (shaded gray region) that extends in time from the removal of the central fixation point to the presentation of the targets. (b) SC activity becomes increasingly predictive of whether a saccade will be directed towards the neuron's preferred (black) or unpreferred (red) target as the time of target presentation approaches. (c) and (d) Testing the functionality of this biased SC activity for preparing saccades. (c) On most trials, the monkey directs a saccade to one of the two target stimuli (crosses); occasionally, SC stimulation triggers a saccade before the targets are presented and to a location orthogonal to the targets (circles). Stimulation-induced saccades deviate slightly towards the ultimately chosen target. (d) Like the neuronal activity recorded in (b), the angular deviation of stimulation-induced saccades increases as the time of target presentation approaches. Each data point represents the mean and standard error of the mean from 6 SC stimulation sites. D. Thevarajah, R. Webb, and M. C. Dorris, unpublished observations.

GAMES IN HUMANS

Research Methods

The neural correlates of social decision making in humans have been investigated thus far using a variety of methods. One approach uses functional neuroimaging, namely functional magnetic resonance imaging (fMRI) or positron emission tomography (PET), to image changes in blood flow while subjects are playing interactive games in the MRI or PET scanner, respectively. Subjects view computer-projected visual stimuli from inside the scanner, either via goggles that display the visual stimuli or via a mirror that allows the subject to view a projection screen outside the scanner. Because verbal responses can create motion artifacts, subjects generally indicate choices by pressing specific buttons on a response box. With these imaging methods, it is possible to examine regional blood flow during the decision-making epochs of the task and to link these to specific choices.

Imaging studies of social interactions have emerged relatively recently within cognitive neuroscience. Many early fMRI studies presented subjects with stimuli of other human faces, given the obvious importance of faces in human social interactions. Typically, these stimuli were static, two-dimensional pictures of faces that subjects were instructed to either passively view, or judge on some attribute such as gender or age (see, for example, Winston *et al.*, 2002). Using similar types of stimuli, others have attempted to probe social cognition by asking subjects to read stories or view cartoons and then make judgments about these hypothetical scenarios. For example, the neural correlates of both mentalizing (Gallagher *et al.*, 2000) and moral reasoning (Greene *et al.*, 2001) have been probed with this methodology. These studies have yielded valuable insights with respect to the neural underpinnings of human social cognition. However, for each, questions can be raised regarding the ecological validity of the stimuli. Does the pattern of brain activation in response to the picture of a static, two-dimensional face accurately reflect the brain's response to the dynamic, embodied faces that we encounter in everyday life? Is the pattern of brain activation in response to reasoning regarding hypothetical, fictitious scenarios the same as when grappling with significant real-life social problems? Is mentalizing about the actions of another person the same as making a consequential decision based on these actions?

One approach to improving the ecological validity of experiments in neuroeconomics is to image brain function as subjects actually interact with other people in real social exchanges (see, for example, McCabe *et al.*, 2001). Recent innovative studies have imaged human subjects while playing both trust and bargaining games with partners communicating from outside the scanner. Potentially even more exciting, hyperscanning technology has been developed that makes it possible to image brain function in two or more interacting partners simultaneously, by utilizing network connections between two separate scanners (e.g. Montague *et al.*, 2002). Hyperscanning has obvious advantages in terms of data collection efficiency (i.e., collecting twice as much data in the same amount of time), but will also open new vistas in social cognitive neuroscience – for example, it will allow imaging of coordinated patterns of brain activity in people who are effectively working together towards a common goal. Further applications for this method will undoubtedly emerge in the future.

Another approach to investigating the neural correlates of social decision making involves manipulating specific neurotransmitter systems and examining the effect on game-playing behavior. For example, dietary tryptophan depletion can be used to decrease brain serotonin levels with a corresponding decrease in cooperative behavior (e.g. Wood *et al.*, 2006), and central oxytocin (OT) levels can be elevated by intranasal self-administration of OT with a corresponding increase in trust (Kosfeld *et al.*, 2005; see also Chapter 15 in this volume). Still another approach involves the use of transcranial magnetic stimulation (TMS) to temporarily activate or deactivate a brain region and then examine its effects on decision making (e.g. van 't Wout *et al.*, 2005). Finally, patients with circumscribed brain damage to particular regions can be tested in these games to see if the damaged brain area has an impact on social decision making (e.g. Koenigs and Tranel, 2007).

Current Research Directions

Use of these innovative methods has allowed researchers to begin to assess brain function as players interact with one another while playing economic games with real consequences. These games have already helped to illuminate facets of the decision-making process – in particular, the degree to which social motives are important in ostensibly economic decisions, and also the processes that may underlie demonstrations of cooperation and competition.

Social Motivation Reward

Decision Making often takes place between options that may be delivered in different modalities – for example, when we are offered the choice between a week's vacation or an extra pay check. Therefore, a common reward mechanism is a crucial component of this system, and a large focus of the broader neuroeconomic endeavor in recent years has been to illuminate the neural processes involved in the encoding and representation of reward, and how these mechanisms may in turn underlie standard models of economic choice such as utility theory and its variants. Part 3 of this volume covers this topic in detail; hence this section will only briefly review the research in this area that pertains to the use of games.

One strong candidate for reward-encoding metric is the mesencephalic dopamine system, and indeed single cell recordings from dopamine neurons and neurons in the striatum, a major projection site of midbrain dopamine cells (see Figure 6.5) have shown that neural responses scale reliably with reward magnitude (Cromwell and Schultz, 2003). Functional neuroimaging studies have corroborated these findings, with studies revealing activation in these areas corresponding with the receipt of reward. Changes in the activity of the striatum have been shown to scale directly with the magnitude of monetary reward or punishment (O'Doherty, 2004; Knutson and Cooper, 2005).

An important development in the investigation of decision making in games has been the discovery that the human striatum appears to also play a central role in social decisions. Importantly, activation of the

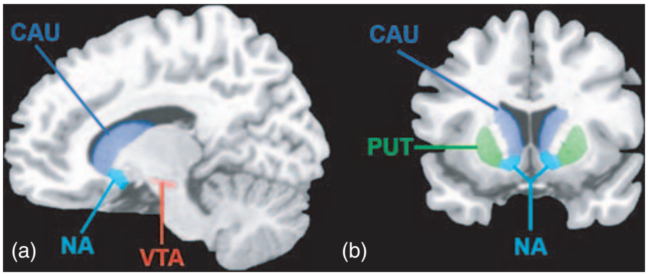


FIGURE 6.5 Brain areas involved in the encoding of reward. Sagittal section (a) and coronal section (b) show the location of the caudate (CAU), nucleus accumbens (NA), and ventral tegmental area (VTA).

striatum in conjunction with social decision making appears to occur above and beyond the financial outcome that may accrue to the player. As will be outlined below, several neuroimaging studies have demonstrated that the striatum tracks a social partner's decision to reciprocate or not to reciprocate cooperation in TG and PDG. One interpretation of these findings is that this area may also encode more abstract valuations, such as the positive feeling garnered by mutual cooperation or the negative feeling of being treated poorly.

For example, reciprocated cooperation with another human in a PDG leads to increased activation in both caudate and nucleus accumbens, as compared to a control condition where an identical amount of money is earned without social involvement. Conversely, unreciprocated cooperation – that is, you cooperate while your partner does not – shows a corresponding decrease in activation in this area (Rilling *et al*., 2002). Additionally, the striatum may be utilized as a guide to informing future decisions in an iterated version of this game, where you must play multiple rounds with the same partner. In these situations, striatal activation on a given round is associated with increased cooperation in subsequent rounds, suggesting that the striatum may register social prediction errors to guide decisions about reciprocity.

Similar findings have been reported in a multiround TG (King-Casas *et al*., 2005). In this version of the TG, participants play several sequential trust games with the same partner, a design that allows examination of how trust is created and signaled within the context of a two-player interaction. In this study, activation in the trustee's caudate was related to how much reciprocity the investor had shown on previous trials, thus corresponding to an "intention to trust" signal of the trustee. Further, this signal gradually shifted in time; in early trials the signal occurred after the investor made her choice, whereas later on this signal occurred much earlier – before, in fact, the investor's decision was revealed.

Of course, social reward need not always be related to positive, mutually cooperative actions. Players also may derive satisfaction from punishing defectors for their norm violations, even when this punishment entails a financial loss to the player. This is illustrated by a PET study (de Quervain *et al*., 2004) in which investors were confronted with non-reciprocators in a TG – that is, players who opted not to return any of the transferred amount. Players had the option to punish these partners by reducing their payout, though, importantly, this action also entailed a loss of points for themselves. Nonetheless, players made the decision to "altruistically punish" in many cases. These decisions were associated with activation in the caudate nucleus, with this activation greater when the punishment was real (involving a financial loss to the other player) than when it was merely symbolic.

Though these rather basic reward and punishment mechanisms have the potential to strongly guide behavior even in complex social decision-making tasks, these prediction error signals can be greatly modulated by top-down processes, such as declarative information previously learned about a partner. For example, in another recent TG study (Delgado *et al*., 2005), players were provided with brief personality sketches of their partners prior to game play. Some partners were described in morally positive terms (for example, by noting how they had recently rescued a person from a fire) and some partners were described in morally negative terms (by describing an unsavory act they had committed). Results demonstrated reduced caudate activity in response to actions of the morally positive or negative partners, though responses to morally neutral players remained unchanged. This suggests that prior social knowledge about a partner can reduce the amount of trial-by-trial learning, demonstrating both top-down and bottom-up influences on the neural basis of social cooperation.

Finally, two recent studies have examined the neural basis of social altruism, by assessing neural activation in tasks where players must decide whether to donate money to charitable organizations. In one study (Moll *et al*., 2006), the striatum was engaged both by receiving money and by donations to charity. Further, the magnitude of this activation was related positively to the number of decisions to donate made by players. In another (Harbaugh *et al*., 2007), these areas were also activated by receipt of money and observing a donation to a charity, but this activation was enhanced when this charitable donation was voluntary as opposed to forced. These latter studies are intriguing, and offer the possibility of extending investigations of social reward beyond simple two-player interactions to questions regarding interactive decision making at a societal level, and potentially have important implications for informing public policy.

It is important to note that some degree of caution should be used when attempting to "reverse-engineer" the interpretation of cognitive and social processes from patterns of brain activity. For example, the association of a brain region with value encoding in previous studies does not necessarily mean that activation in this area in the context of an interactive game can automatically be interpreted as rewarding or punishing. It would therefore be prudent for the field as a whole to buttress these claims by either converging evidence from other methodologies such as TMS or patient work, or at the very least demonstrating behavioral performance in line with the neural predictions, such as a player's preference for options that activate reward centers more strongly (e.g. de Quervain *et al*., 2004).

Nonetheless, these results do appear to demonstrate that complex social processes recruit more basic mechanisms within the human brain, providing support for the notion that the brain uses a common reward metric, and also informing economic theories of reciprocity and inequity aversion (e.g. Dufwenberg and Kirchsteiger, 2004). This also furthers the connection between the disparate branches of neuroeconomics, as it suggests that research into the processing of primary and secondary rewards (such as food and money) may be directly applicable to how we encode more abstract social rewards like reciprocity and fairness.

Competition, Cooperation and Coordination

Use of games that evoke often quite powerful feelings of competitiveness or camaraderie have helped to illuminate the complex nature of processing that occurs while engaged in a social decision-making situation. In addition to the rewarding or punishing effects of social interactions, as exemplified by neural activation in classical "reward" brain regions described above, these interactive scenarios have also illustrated the prominent role emotions play in decision-making games. Classical models of decision making, both utility theory for individual decisions and game theory for social decisions, have largely ignored the influence of emotions on how decisions are made, but recent research has begun to demonstrate their powerful effect.

Emotional processes seem to reliably engage a set of brain structures including reward-processing mechanisms discussed above, areas of the midbrain and cortex to which they project (such as ventromedial frontal cortex (VMPFC), orbitofrontal (OFC), and anterior cingulate cortex (ACC)), as well as a number of other areas such as the amygdala and insula (Dalgleish, 2004) (see Figure 6.6).

Early pioneering work in this domain showed that patients suffering damage to VMPFC, who presented with associated emotional deficits, were impaired on gambling tasks (Damasio, 1994; Bechara and Damasio, 2005), demonstrating experimentally that emotion plays a vital role in determining decisions. Further research in the behavioral domain (Mellers *et al*., 1999) as well as with functional neuroimaging (Coricelli *et al*., 2005) has shown the biasing effect of emotions such as anticipated regret and disappointment on decision making, specifically demonstrating that people steer clear of potential outcomes that they predict could cause feelings of regret, even if these options have a higher monetary expected value.

In terms of decision making in the context of games, negative emotional states have been observed behaviorally as a result of both inequity and non-reciprocity, such as unfair offers in a UG (Pillutla and Murnaghan, 1996). These emotional reactions have been proposed as a mechanism by which inequity is avoided, and may have evolved precisely to foster mutual reciprocity, to make reputation important, and to punish those seeking to take advantage of others (Nowak *et al*., 2000). Indeed, even capuchin monkeys respond negatively to unequal distributions of rewards by refusing to participate in an effortful task if they witness another player receiving equal reward for less work (Brosnan and de Waal, 2003 – see also Chapter 19 of this volume).

Neuroscientific studies of this nature offer the potential to go beyond speculation and to examine the causal relationship between an emotional reaction and subsequent social decision, as well as investigating whether areas specialized for the processing of basic emotions may be coopted for more complex affective reactions.

To examine this question, and more broadly to attempt to better specify the systems involved in the neurobiology of social decision making, Sanfey *et al*. (2003) conducted a neuroimaging study examining the brain's response to fair and unfair offers in a UG, and in particular to investigate how these responses were related to the decision to accept or reject in the game. Participants were scanned using fMRI as they played the role of responder in the UG. Prior to scanning, each participant was introduced to 10 people they were told would partner with them in the game. The offers that the participants saw were in fact predetermined, with half being fair (a \$5:\$5 split of a \$10 pot) and half being unfair (two offers of \$9:\$1, two offers of \$8:\$2, and one offer of \$7:\$3). This distribution of offers generally mimics the range of offers typically

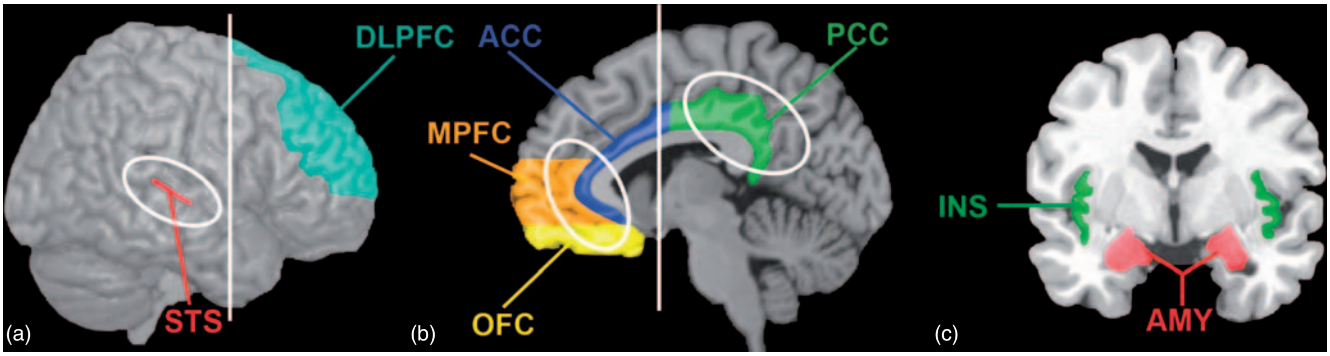


FIGURE 6.6 Map of brain areas commonly found to be activated in social decision-making studies. The lateral view (a) shows the location of the dorsolateral prefrontal cortex (DLPFC) and superior temporal sulcus (STS). The sagittal section (b) shows the location of the anterior cingulate cortex (ACC), medial prefrontal cortex (MPFC), orbitofrontal cortex (OFC), and posterior cingulate cortex (PCC). The coronal section (C, cut along the white lines in both A and B) shows the location of the insula (INS) and amygdala (AMY). Areas circled are those often associated with Theory of Mind processes.

made in uncontrolled versions of the game (i.e. involving freely acting human partners). Players also saw 10 offers from a computer partner identical to those from the human partners, which were introduced to distinguish between intentional offers made by other players and the same offers made by a random device.

Behavioral results in this experiment were very similar to those typically found in UG studies. Participants accepted all fair offers, with decreasing acceptance rates as the offers became less fair. Unfair offers of $2 and $1 made by human partners were rejected at a significantly higher rate than the same offers made by a computer, suggesting that participants had a stronger emotional reaction to unfair offers from humans than to those from computers.

With regard to neuroimaging, the contrast of primary interest was between the neural responses to unfair offers as compared to fair offers. The brain areas showing greatest activation for this comparison were bilateral anterior insula, dorsolateral prefrontal cortex (dlPFC), and anterior cingulate cortex (ACC). In bilateral insula, the magnitude of activation was also significantly greater for unfair offers from human partners as compared to both unfair offers from computer partners and control amounts, suggesting that these activations were not solely a function of the amount of money offered to the participant but were also uniquely sensitive to the context – namely, perceived unfair treatment from a human. Also, regions of bilateral anterior insula demonstrated sensitivity to the degree of unfairness of an offer, exhibiting significantly greater activation for a $9:$1 offer than an $8:2 offer from a human partner.

Activation of anterior insula to unfair offers from human partners is particularly interesting in light of this region's oft-noted association with negative emotional states (Derbyshire *et al.*, 1997; Calder *et al.*, 2001). Anterior insula activation is consistently seen in neuroimaging studies of pain and distress, of hunger and thirst (Denton *et al.*, 1999), and of autonomic arousal (Critchley *et al.*, 2000). Further, right anterior insula activity has been implicated in aversive conditioning (Seymour *et al.*, 2005). In a related study, this area was also active in an iterated prisoner's dilemma game (Rilling *et al.*, 2008), where individuals with a stronger anterior insula response to unreciprocated cooperation showed a higher frequency of defection. These results suggest that anterior insula and associated emotion-processing areas may play a role in marking a social interaction as aversive, and thus discouraging trust of the partner in the future.

If the activation in the anterior insula is a reflection of the responders' negative emotional response to an unfair offer, we might expect activity in this region to correlate with the subsequent decision to either accept or reject the offer. Indeed, collapsing across participants, an examination of individual trials revealed a relationship between right anterior insula activity and the decision to accept or reject; namely, that a higher insula response to an unfair offer was related to higher rejection rates of these offers. Separate measures of emotional arousal provide support for this hypothesis. A UG study measuring skin-conductance responses, used as an autonomic index of affective state, found that the skin conductance activity was higher for unfair offers, and, as with insula activation, discriminated between acceptances and rejections of these offers (van 't Wout *et al.*, 2006).

In contrast to the insula, dlPFC usually has been linked to cognitive processes such as goal maintenance and executive control. In a similar vein to the suppression of striatal activation by frontal, "top-down" processes in reward studies, we can interpret the activation of frontal regions to unfair offers in UG studies as a mechanism by which other more deliberative goals (such as reputation maintenance or the desire to make money) can be implemented.

Of course, as with all brain-imaging data, these results are largely correlative, but they do provide hypotheses for further testing – namely, that activation of areas associated with emotion processing (in this case the anterior insula) is related to the negative experience of receiving an unfair offer from another human, and as such is related to the decision to reject, while activation of frontal, more traditionally deliberative regions such as dlPFC may represent the cognitive goals of the task. Therefore, a further set of studies has sought to target these brain areas with a variety of methods in order to examine whether accept/reject decisions in the UG could be manipulated via these purported mechanisms.

As mentioned above, activation of frontal regions to unfair offers in UG studies has been interpreted as a mechanism by which other more deliberative goals (such as reputation maintenance or the desire to make money) can be implemented. In two novel studies (van 't Wout *et al.*, 2005; Knoch *et al.*, 2006), transcranial magnetic stimulation (TMS) was used to disrupt processing in dlPFC while players were making decisions about offers in a UG. In both cases, stimulation increased acceptance rate of unfair offers as compared to control, providing strong evidence for a causal relationship between activation in this area and social decision making. Though TMS is still rather a crude tool (thus making clear-cut interpretations of behavior challenging), use of this technology, as well as behavioral and other neuroimaging work, to experimentally test hypotheses generated by this early series of studies will be vital in progressing the field.

In concert with this investigation of the deliberative system, experimental methods have also been used to prime the affective system. The initial fMRI UG experiment described above demonstrated that the decision to reject offers in the UG is strongly correlated with increases in activation of the anterior insula. To directly investigate the relationship between negative emotional states, activation of the anterior insula, and decisions to reject unfair offers, a follow-up experiment was conducted in which negative emotion states were primed prior to playing the UG (Harle and Sanfey, 2007). The hypothesis to be tested here was that the priming of negative emotion states known to engage the anterior insula, such as sadness and disgust (Damasio *et al.*, 2000), would lead to higher rejection rates of unfair offers.

Prior to playing as responder in the standard UG, participants in this study viewed a 5-minute video that was ostensibly unrelated to the UG section of the experiment. These clips had been previously rated as "sad," "happy," or "neutral" by a separate group of participants. The primary research finding was that the group of participants who viewed the "sad" video (an excerpt from the movie *The Champ*) had an overall significantly higher rejection rate of unfair offers than those who watched either the neutral or the happy clip, indicating a demonstrable effect of negative mood on "emotional" decisions in the UG. This is important, as it shows that subtle and transient emotional states, unrelated to the task at hand, can noticeably affect decisions to accept or reject monetary offers. Further, it suggests a causal relation between negative emotional states, activation of specific affectively specialized brain regions, such as anterior insula, and decision making. It also suggests that examining decision-making performance in participants with dis-regulated emotional processing, such as patients with depression or schizophrenia, may be a useful future avenue of research. Indeed, patients with damage to ventromedial prefrontal cortex, another area implicated in the processing of emotional information, also reject unfair offers more frequently than do controls (Koenigs and Tranel, 2007).

The findings outlined above provide an initial toehold for measuring physical mechanisms responsible for social decision making in the brain. Such studies offer the promise that we will be able to identify and precisely characterize these mechanisms, and the factors that influence their engagement and interaction. Even at this early stage, however, results highlight the fact that decision making appears to involve the interaction among multiple subsystems governed by different parameters and possibly even different principles.

Finally, while the research reviewed here has greatly increased our understanding of the neural correlates of social decisions, it is important to note that these data also have the potential to inform economic theories of interactive decisions-making. Recent models in behavioral economics have attempted to account for social factors, such as inequity aversion, by adding these social utility functions to the standard models (see, for example, Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000), and modeling these functions based on the underlying neural patterns may provide useful constraint on these models.

CONCLUSION

The preceding sections review some general ways in which the tasks of experimental economics and the techniques of neuroscience can make important contributions to the understanding of social decision making. In particular, we have focused on human and non-human techniques that provide insight into the physical mechanisms responsible for decision making in games.

The invasive techniques used in non-human primates allow neural activity to be recorded at high spatial and temporal resolution, and correlated to specific stages of game play or behavior. Furthermore, the functionality of localized patterns of neural activities on game play can be examined through artificial manipulation. Ultimately, however, we are more concerned with the human condition. Therefore, neural mechanisms must be studied while humans are engaged in social interactions. Fortunately, for each invasive experimental technique employed in non-human primates, analogous (if somewhat less sensitive) non-invasive techniques exist for studying brain processes in humans. For example, single neuron activity is complemented by non-invasive brain-imaging methods such as functional magnetic resonance imaging, positron emission tomography, and electroencephalography. Local perturbation of neural tissue with micro-stimulation is complemented by transcranial magnetic stimulation. Reversible manipulation of neural activity with pharmacological agents is complemented with studies involving patients with focal brain damage and those receiving systemic application of pharmacological agents.

Of course, progress in our understanding will occur most rapidly when both invasive and non-invasive techniques can be brought to bear on the same underlying processes. For example, as outlined above, both human and non-human primates display similar strategies during simple mixed-strategy games and matching-law tasks. Patterns of behavioral choice and neuronal activity are well described by reinforcement learning algorithms under these conditions, suggesting a foundational role for this class of learning during social interactions.

Progress in delineating more sophisticated cognitive modules during social interactions will rely heavily on our ability to design appropriate laboratory tasks. These tasks must be amenable to state-of-the-art neuroscience techniques, and yet still capture the essence of natural social interactions. Human experiments must overcome the isolation that typically accompanies current brain-imaging technologies. Great strides have been made in this area, by providing more realistic and interactive displays, and with the use of hyperscanning technology. It is particularly difficult to assess whether non-human primates realize they are involved in social interactions, because of course they cannot receive verbal instructions nor self-report on their experiences. For example, in the experiments conducted to date, animals may perceive that they are simply chasing dots of light that occasionally elicit rewards rather than being involved in an interactive game. If the claim is made that particular neural activities subserve social interactions rather than more straightforward reward mechanisms, future non-human primate studies should incorporate interactions between real (or perhaps virtual?) cohorts during game play.

Appropriate task design is particularly important for distinguishing between those cognitive modules that are present to a lesser degree in non-human primates from those that are posited as completely unique to humans. For example, it has been long debated whether a theory of mind module, which is critical for inferring the beliefs and intentions of others during game play, exists in non-human primates (Penn and Povinelli, 2007). Similarly, in certain contexts non-human primates display what appears to be an innate sense of fairness (Brosnan and De Waal, 2003). However, when tested on a laboratory version of the ultimatum game, non-human primates' strategies did not reflect the unfair play of their opponent (Jensen *et al*., 2007). These examples illustrate the difficulty in determining whether non-human primates lack particular cognitive modules or whether current laboratory tasks fail to capture the socially or ecologically relevant aspects of natural environments.

In conclusion, this chapter has examined recent attempts at combining neuroscience methodology with the theoretical framework of game theory. Recent research illustrates the potential for this cross-disciplinary approach to provide advances in our understanding of the neural mechanisms subserving sophisticated social interactions.

References

Andersen, R.A. (1995). Encoding of intention and spatial location in the posterior parietal cortex. *Cerebral Cortex* 5, 457–469.

Balleine, B.W., Delgado, M.R., and Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *J. Neurosci.* 27, 8161–8165.

Barraclough, D.J., Conroy, M.L., and Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* 7, 404–410.

Bechara, A. and Damasio, A.R. (2005). The somatic marker hypothesis: a neural theory of economic decision. *Games Econ. Behav.* 52, 336–372.

Bisley, J.W. and Goldberg, M.E. (2003). Neuronal activity in the lateral intraparietal area and spatial attention. *Science* 299, 81–86.

Bolton, G.E. and Ockenfels, A. (2000). ERC: a theory of equity, reciprocity, and competition. *Am. Econ. Rev.* 90, 166–193.

Brosnan, S.F. and de Waal, F.B.M. (2003). Monkeys reject unequal pay. *Nature* 425, 297–299.

Calder, A.J., Lawrence, A.D., and Young, A.W. (2001). Neuropsychology of fear and loathing. *Nat. Rev. Neurosci.* 2, 352–363.

Camerer, C.F. (2003). *Behavioral Game Theory*. Princeton, NJ: Princeton University Press.

Carello, C.D. and Krauzlis, R.J. (2004). Manipulating intent: evidence for a causal role of the superior colliculus in target selection. *Neuron* 43, 575–583.

Coricelli, G., Critchley, H.D., Joffily, M. *et al*. (2005). Regret and its avoidance: a neuroimaging study of behavior. *Nat. Neurosci.* 8, 1255–1262.

Corrado, G.S., Sugrue, L.P., Seung, H.S., and Newsome, W.T. (2005). Linear–nonlinear-Poisson models of primate choice dynamics. *J. Exp. Anal. Behav.* 84, 581–617.

Critchley, H.D., Elliott, R., Mathias, C.J., and Dolan, R.J. (2000). Neural activity relating to generation and representation of galvanic skin conductance responses: a functional magnetic resonance imaging study. *J. Neurosci.* 20, 3033–3040.

Cromwell, H.C. and Schultz, W. (2003). Influence of the expectation for different reward magnitudes on behavior-related activity in primate striatum. *J. Neurophysiol.* 89, 2823–2838.

Dalgleish, T. (2004). The emotional brain. *Nat. Rev. Neurosci.* 5, 583–589.

Damasio, A.R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York, NY: Putnam.

Damasio, A.R., Grabowski, T.J., Bechara, A. *et al*. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nat. Neurosci.* 3, 1049–1056.

Delgado, M.R., Frank, R.H., and Phelps, E.A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8, 1611–1618.

Denton, D., Shade, R., Zamarippa, F. *et al*. (1999). Neuroimaging of genesis and satiation of thirst and an interoceptor-driven theory of origins of primary consciousness. *Proc. Natl Acad. Sci. USA* 96, 5304–5309.

de Quervain, D.J., Fischbacker, U., Treyer, V. *et al*. (2004). The neural basis of altruistic punishment. *Science* 305, 1254–1258.

Derbyshire, S.W., Jones, A.K.P., and Gyulai, F. (1997). Pain processing during three levels of noxious stimulation produces differential patterns of central activity. *Pain* 73, 431–445.

Dorris, M.C. and Glimcher, P.W. (2004). Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44, 365–378.

Dorris, M.C., Pare, M., and Munoz, D.P. (1997). Neuronal activity in monkey superior colliculus related to the initiation of saccadic eye movements. *J. Neurosci.* 17, 8566–8579.

Dorris, M.C., Olivier, E., and Munoz, D.P. (2007). Competitive integration of visual and preparatory signals in the superior colliculus during saccadic programming. *J. Neurosci.* 27, 5053–5062.

Dufwenberg, M. and Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games Econ. Behav.* 47, 268–298.

Fehr, E. and Schmidt, K.M. (1999). A theory of fairness, competition, and cooperation. *Q. J. Economics* 114, 817–868.

Gallagher, H.L., Happe, F., Brunswick, N. *et al*. (2000). Reading the mind in cartoons and stories: an fMRI study of "theory of mind" in verbal and nonverbal tasks. *Neuropsychologia* 38, 11–21.

Glimcher, P.W. (2003). The neurobiology of visual-saccadic decision making. *Annu. Rev. Neurosci.* 26, 133–179.

Glimcher, P.W. and Sparks, D.L. (1992). Movement selection in advance of action in the superior colliculus. *Nature* 355, 542–545.

Gold, J.I. and Shadlen, M.N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature* 404, 390–394.

Gold, J.I. and Shadlen, M.N. (2003). The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *J. Neurosci.* 23, 632–651.

Goldberg, M.E., Bisley, J.W., Powell, K.D., and Gottlieb, J. (2006). Saccades, salience and attention: the role of the lateral intraparietal area in visual behavior. *Prog. Brain Res.* 155, 157–175.

Grantyn, A., Moschovakis, A.K., and Kitama, T. (2004). Control of orienting movements: role of multiple tectal projections to the lower brainstem. *Prog. Brain. Res.* 143, 423–438.

Greene, J.D., Sommerville, R.B., Nystrom, L.E. *et al*. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105–2108.

Grefkes, C. and Fink, G.R. (2005). The functional organization of the intraparietal sulcus in humans and monkeys. *J. Anatomy* 207, 3–17.

Guth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Org.* 3, 376–388.

Harbaugh, W.T., Mayr, U., and Burghart, D.R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* 316, 1622–1625.

Harle, K. and Sanfey, A.G. (2007). Sadness biases social economic decisions in the Ultimatum Game. *Emotion* 7, 876–881.

Harsanyi, J.C. (1974). Equilibrium-point interpretation of stable sets and a proposed alternative definition. *Manag. Sci. Series A – Theory* 20, 1472–1495.

Henrich, J., Boyd, R., Bowles, S. *et al*. (2005). Economic man' in cross-cultural perspective: ethnography and experiments from 15 small-scale societies. *Behav. Brain Sci.* 28, 795–855.

Jensen, K., Call, J., and Tomasello, M. (2007). Chimpanzees are rational maximizers in an ultimatum game. *Science* 318, 107–109.

Kahneman, D., Slovic, P., and Tversky, A. (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Kalaska, J.F., Cisek, P., and Gosselin-Kessiby, N. (2003). Mechanisms of selection and guidance of reaching movements in the parietal lobe. *Adv. Neurol.* 93, 97–119.

King-Casas, B., Tomlin, D., Anen, C. *et al*. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83.

Knoch, D., Pascual-Leone, A., Meyer, K. *et al*. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314, 829–832.

Knutson, B. and Cooper, J.C. (2005). Functional magnetic resonance imaging of reward prediction. *Curr. Opin. Neurol.* 18, 411–417.

Koenigs, M. and Tranel, D. (2007). Irrational economic decision-making after ventromedial prefrontal damage: evidence from the Ultimatum Game. *J. Neurosci.* 27, 951–956.

Kosfeld, M., Heinrichs, M., Zak, P.J. *et al*. (2005). Trust in a bottle. *Nature* 435, 673–676.

Lau, B. and Glimcher, P.W. (2005). Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* 84, 555–579.

Lee, C., Rohrer, W.H., and Sparks, D.L. (1988). Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature* 332, 357–360.

Lee, D., Conroy, M.L., McGreevy, B.P., and Barraclough, D.J. (2004). Reinforcement learning and decision making in monkeys during a competitive game. *Brain Res. Cogn. Brain Res.* 22, 45–58.

Lee, D., McGreevy, B.P., and Barraclough, D.J. (2005). Learning and decision making in monkeys during a rock–paper–scissors game. *Brain Res. Cogn. Brain Res.* 25, 416–430.

McCabe, K., Houser, D., Ryan, L. *et al*. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl Acad. Sci. USA,* 98, 11832–11835.

Mellers, B., Schwartz, A., and Ritov, I. (1999). Predicting choices from emotions. *J. Exp. Psychol. Gen.* 128, 332–345.

Moll, J., Krueger, F., and Zahn, R. *et al*. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc. Natl Acad. Sci. USA*, 103, 15623–15628.

Montague, P.R., Berns, G.S., Cohen, J.D. *et al*. (2002). Hyperscanning: simultaneous fMRI during linked social interactions. *NeuroImage* 16, 1159–1164.

Nash, J.F. (1950). Equilibrium points in n-person games. *Proc. Natl Acad. Sci. USA* 36, 48–49.

Nowak, M.A., Page, K.M., and Sigmund, K. (2000). Fairness versus reason in the Ultimatum Game. *Science* 289, 1773–1775.

O'Doherty, J.P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* 14, 769–776.

Pare, M. and Wurtz, R.H. (2001). Progression in neuronal processing for saccadic eye movements from parietal cortex area lip to superior colliculus. *J. Neurophysiol.* 85, 2545–2562.

Penn, D.C. and Povinelli, D.J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a "theory of mind". *Phil. Trans. R. Soc. Lond. B Biol.* 362, 731–744.

Pillutla, M.M. and Murnighan, J.K. (1996). Unfairness, anger, and spite: emotional rejections of ultimatum offers. *Org. Behav. Human Dec. Proc.* 68, 208–224.

Platt, M.L. and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.

Rilling, J.K., Gutman, D.A., Zeh, T.R. *et al*. (2002). A neural basis for social cooperation. *Neuron* 35, 395–405.

Rilling, J.K., Goldsmith, D.R., Glenn, A.L. *et al*. (2008). The neural correlates of the affective response to unreciprocated cooperation. *Neuropsychologia* 46, 1256–1266.

Robinson, D.A. (1972). Eye movements evoked by collicular stimulation in the alert monkey. *Vision Res.* 12, 1795–1808.

Romo, R. and Salinas, E. (2003). Flutter discrimination: neural codes, perception, memory and decision making. *Nat. Rev. Neurosci.* 4, 203–218.

Sally, D. and Hill, E.L. (2006). The development of interpersonal strategy: autism, theory-of-mind, cooperation and fairness. *J. Econ. Psychol.* 27, 73–97.

Salzman, C.D., Britten, K.H., and Newsome, W.T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. *Nature* 346, 174–177.

Sanfey, A.G., Rilling, J.K., Aronson, J.A. *et al*. (2003). The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1758.

Schall, J.D. and Thompson, K.G. (1999). Neural selection and control of visually guided eye movements. *Annu. Rev. Neurosci.* 22, 241–259.

Schultz, W., Tremblay, L., and Hollerman, J.R. (2000). Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral Cortex* 10, 272–284.

Seo, H. and Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* 27, 8366–8377.

Seymour, B., O'Doherty, J.P., Koltzenburg, M. *et al*. (2005). Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat. Neurosci.* 8, 1234–1240.

Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.

Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2005). Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* 6, 363–375.

van 't Wout, M., Kahn, R.S., Sanfey, A.G., and Aleman, A. (2005). rTMS over the right dorsolateral prefrontal cortex affects strategic decision making. *NeuroReport* 16, 1849–1852.

van 't Wout, M., Kahn, R.S., Sanfey, A.G., and Aleman, A. (2006). Affective state and decision-making in the Ultimatum Game. *Exp. Brain Res.* 169, 564–568.

von Neumann, J. and Morgenstern, O. (1947). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Winston, J.S., Strange, B.A., O'Doherty, J., and Dolan, R.J. (2002). Automatic and intentional brain responses during evaluations of trustworthiness of faces. *Nat. Neurosci.* 5, 277–283.

Wood, R.M., Rilling, J.K., Sanfey, A.G. *et al*. (2006). The effects of altering 5-HT activity on the performance of an iterated prisoner's dilemma (PD) game in healthy volunteers. *Neuropsychopharmacology* 31, 1075–1084.

CHAPTER

7

The Evolution of Rational and Irrational Economic Behavior: Evidence and Insight from a Non-human Primate Species

Laurie R. Santos and M. Keith Chen

OUTLINE

INTRODUCTION

Modern economics as it is currently practiced is an exercise in applying three basic principles to nearly all settings. First, it entails positing agents with simple, stable preferences. Workers are assumed to maximize earnings net their disutility of labor, consumers are assumed to maximize a stable utility function given their budgets, and family members are assumed to bargain with each other given their competing goals. Second, people are endowed with effortlessly rational, error-free cognition. This assumption may entail agents simply understanding their own preferences, or it may ask that they solve arbitrarily complex signal-extraction problems. Finally, modern economics assumes that people interact with each other in ways that are relatively frictionless and thus yield equilibrium behavior. That is, people are assumed to maximize their own interests given the behavior of others, equalizing their personal returns across activities.

All three of these assumptions have proven deeply useful to economists. Assuming simple preferences limits the degree to which the analyst might "overfit" behaviors, and stable preferences are necessary if current observations are to bear any predictions about different contexts or future events. Assuming rational agents and equilibrium outcomes likewise disciplines analysts, making sure their predictions depend more on observable facts about the environment than they do on unobservable psychological properties, which are undoubtedly more difficult to measure and quantify. Unfortunately, although assumptions about stable preferences have proven formally useful to economists, it is clear that human decision makers do not always live up to the modern economists' high standards. Behavioral economists have spent the last few decades documenting a number of systematic ways in which human consumers violate standard economic assumptions (see reviews in Camerer, 1998; Kahneman *et al.*, 1982).

Given the systematic errors and biases that psychologists and behavioral economists study, it may at first glance seem foolish to embark on a study of economic behavior and preferences in other species. If humans can't perform fast and error-free computations, achieve equilibrium reliably, or maintain stable and frame-invariant preferences, it seems unlikely that other, presumably less computationally-savvy, species will be able to do so. Nevertheless, this chapter will argue that modern economics – and, importantly, the emerging field of neuroeconomics – can gain insight into the nature of human preferences through the study of other species, particularly other closely related primates. While we agree that the behavior of non-human primates may have little hope of shedding light on such hyper-rational agents and their economies, we will argue that research examining non-human primate preferences may have something important to teach us about the deep structure of human preferences, and the way that less-than-perfect agents with those preferences respond to incentives.

This chapter will review our recent discoveries about preferences in one model primate species – the capuchin monkey. We begin by reviewing a number of different economic approaches to non-standard choice behavior in humans. We will then turn to our own work exploring whether capuchin monkeys (*Cebus apella*) also exhibit non-standard choice behavior in situations analogous to those seen in human markets. We will use this work to argue that many of the central lessons of price theory hold in (presumably) less than fully rational capuchin economies, and that many of the aspects of the prospect-theoretic preferences we observe in humans also appear in capuchin behavior. Observing that non-human primates display the same fundamental biases that humans do, and that these biases respond similarly to incentives, suggests both an expanded role for these biases in positive accounts of human economies, and that these biases may form the basis for a stable set of deeper preferences towards which economic tools can be applied.

NEOCLASSICAL APPROACHES TO NON-STANDARD BEHAVIOR

Although economists often formally assume that humans are hyper-rational agents, most economists recognize that humans commonly fail to live up to the standard of *Homo economicus*. Indeed, neither Adam Smith, the founder of classical economics, nor Alfred Marshall thought that humans were perfectly rational agents, and neither thought that rationality was a necessary condition for the usefulness of price theory. Instead, classical economists hypothesized that agents had and were motivated by simple, stable, self-interested preferences, and that such preferences acted to equalize returns across different activities, eliminating arbitrage opportunities and inducing efficient markets. As Smith famously wrote, "it is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own interest."

Price-Theoretic Treatments

Neoclassical economists realized that their insights did not require agents to be hyper-rational; agents simply needed to respond to incentives. Under this view of agents, then, behavioral biases and cognitive limitations can be fruitfully studied using neoclassical economic techniques. One of the classic examples of this approach is the work of Gary Becker. As Becker (1962) himself put it, "the important theorems of modern economics result from a general principle which not only includes rational behavior and survivor arguments as special cases, but also much irrational behavior." Consistent with this idea, Becker and co-authors have used price-theoretic tools in settings which economists had previously thought not amenable to rational analysis. In the essays collected in his seminal *Economic Approach to Human Behavior* (1976), Becker applies price theory to understand such diverse phenomena as racial discrimination, family dynamics, and crime and punishment. In perhaps the most pure example of this approach, Becker and Murphy (1988) analyzed addictive behavior by positing that such behavior may arise from underlying stable preferences in which consumption of an addictive good today is a complement to consumption of that same good tomorrow. This price-theoretic framework yields important insights into addictive behavior, including rapidly increasing or declining (yet perfectly rational) consumption of addictive goods, "cold-turkey" quitting strategies, and the prediction that addicts will respond much more to permanent than to temporary price changes.

Becker's approach relies on assuming that what might seem transient, unstable, and irrational behavior may actually arise from stable, underlying preferences. These preferences may include terms not normally included in the arguments of utility – terms such as altruism, fairness, tastes, habits, and prejudices. Positing these more basic, stable preferences is fundamental to the application of neoclassical tools to non-standard settings. For example, Becker writes that "generally (among economists) … preferences are assumed not to change substantially over time, nor to be very different between wealthy and poor persons, or even between persons in different societies and cultures." Indeed, coupled with maximizing behavior and market equilibrium, Becker asserts that the assumption of stable preferences "forms the heart of the economic approach."

More recently, Ed Glaeser (2004) has argued that even if researchers were to show that human decision making is driven more by temporary, fleeting, situational factors than it is by stable preferences, this would only serve to increase the importance of classic price-theoretic techniques. This is because "many topics require both psychological insight into the power of local influence and economic reasoning about the supply of that influence" (Glaeser, 2004). Thus, even if it were the case that people made decisions based strongly on temporary and situational cues, in most market situations those cues will be provided by self-interested entrepreneurs such as marketers or politicians. Glaeser argues that price-theory is essentially the only tool we have to understand the supply of such frames and persuasive messages. The payoff to such an approach, Glaeser asserts, is powerful in that predictions arise from an equilibrium analysis of the supply of such messages. For example, Glaeser (2004) notes that:

> The applications of economics to the formation of aggregate cognitive errors suggest a number of comparative statics. These errors will be more common when the costs of making mistakes to the individual are low. As a result, we should expect more errors in the political arena (because no one's vote directly matters) than in the market arena (because making foolish purchases is at least somewhat costly). These errors will be more common when mistaken beliefs strongly complement supplier's returns. Mistaken beliefs will be more common when errors increase the current flow of utility. Thus, if people enjoy anticipating a rosy future, they should believe stories that make them overly optimistic and in particular, they should happily accept stories about a life after death.

Axiomatic Approaches

Another way neoclassical economists have dealt with non-standard behavior is through the use of axiomatic approaches. Where the price-theoretic approach to non-standard behavior focuses more on the role of incentives and market discipline in shaping (possibly non-standard) behavior, the axiomatic approaches focuses on weakening the assumptions underlying utility theory so as to allow the analysis of non-standard behavior. Kreps and Porteus (1978) used a classic axiomatic approach to study agents who appear to prefer earlier resolution of uncertainty rather than later (or *vice versa*), even though the timing of the resolution has no consequential effects. The Kreps-Porteus approach deals with this temporal inconsistency by applying the classic axioms of choice under uncertainty to dated lotteries – lotteries that specify not just what information will be revealed, but when that uncertainty will be revealed. Kreps-Porteus establishes a representation result that allowed for the prices definition of preferences for early resolution of uncertainty, allowing standard tools of economics to be applied to markets where the timing of information revelation is key, with broad applications in macroeconomics and finance.

More recently, Gul and Pesendorfer applied axiomatic choice theory to the phenomena of dynamic inconsistency and temptation preferences, with hyperbolic discounting being the most widely studied example. For instance, Gul and Pesendorfer (2001) used classic choice theory to study *choice sets*, rather than choices *per se*. A decision maker might, for example, strictly prefer the choice set B to the choice set A, even if A offers strictly more options (B is strictly a subset of A), because some of those options in A might produce temptation costs. Similar to the Kreps-Porteus approach, Gul and Pesendorfer derived a set of axioms which many simple forms of temptation satisfy and showed that, under those axioms, a simple representation of preferences in terms of linear functions suffices. This allows for the rigorous definition and study of markets in which temptation and a demand for self-control may exist. Fundamental to both axiomatic and price-theoretic approaches, however, is a strict neoclassical emphasis on positive economics; alternative axioms and utility functions are to be judged solely by their parsimony and ability to predict choice behavior. Most notably, this de-emphasizes any appeal to psychological realism, one of the main distinctions between neoclassical and behavioral economics.

Behavioral Economics Approaches

In contrast to these neoclassical approaches, much of modern behavioral economics starts by scanning the nearby disciplines of social psychology and sociology for robust biases that may manifest themselves in economically important settings. Economists using this approach have tried to incorporate psychological and sociological findings into economic analysis by finding a functional form for preferences that captures many of the stylized facts that these biases present. Most prominently, Kahneman and Tversky (1979) attempted to unify several stylized deviations from expected utility in a single theory of choice under uncertainty called *prospect theory*. Prospect theory represents choice as a function of the value of the choices rather than as a function of a person's overall utility. These values are assessed as either gains or losses (i.e., positive or negative differences) relative to an arbitrary reference point. A major implication of prospect theory, then, is that decision makers naturally frame their decisions as gains or losses relative to a particular reference point. Prospect theory's value function passes through the reference point as S-shaped, with a kink in the curve at the reference point, such that a given absolute-sized loss (e.g. a $5 loss) will decrease value more than an identically-sized gain (e.g. a $5 gain) will increase value. This feature of the value curve leads to loss-aversion: decision makers are more sensitive to a loss than they are to an equally-sized gain, which can lead to odd and often irrational framing effects in which decision-makers' responses may vary with how the choice is presented, worded, or described. The structure of the value curve also leads to a phenomenon known as the *reflection effect*: decision makers treat changes from a reference point differently depending on whether they are gains or losses. More specifically, decision makers tend to be risk-seeking when dealing with perceived losses, but risk-averse when dealing with perceived gains.

Prospect theory has been widely applied across numerous fields in economics, including finance (explaining the disposition effect and the equity premium), labor supply (income targeting), and consumer choice (asymmetric price elasticities, the endowment effect). (See Camerer (1998) for an elegant and comprehensive review of the applications of prospect theory in economics.)

Another widely used model in behavioral economics is David Laibson's model of time-inconsistent choice. Laibson (1997) modeled inter-temporal inconstancy with a beta-delta model of hyperbolic discounting, and demonstrated how agents with such preferences could be imbedded in economic models of choice over time. By doing this and demonstrating how to solve such the dynamic-programming problem that these agents face when trying to optimize, economists could model the effects of present-biased preferences and how they might interact with different types of illiquid assets, market structures, or public policies.

THE ROLE OF NON-HUMAN PRIMATE STUDIES IN MODERN ECONOMICS

Common to all the approaches reviewed is that, by and large, they take the origins and structure of behavioral biases as given. To date, far less direct attention has been paid to understanding how basic or fundamental these biases are. Put differently, most of the approaches reviewed above explicitly model the external market forces and technologies which shape the supply of cues, yet the cognitive systems and constraints that lead to these biases are worked around, often in one of two ways. Most behavioral economists leave these biases to social psychologists to study, acting essentially as importers of psychological insights. In turn, the models that behavioral economists use are based on assumptions judged not only by their ability to organize economic data, but also by their psychological realism. Axiomatic approaches, in contrast, tend to disregard the latter of these two goals, instead treating the minds of people as black boxes that are approachable through observing choice data alone. In both behavioral economic and axiomatic approaches, however, little work has examined how our behavioral biases arise in the first place.

What, then, are the origins and deeper structure of our systematic economic biases? Are our biases the result of social or cultural learning and specific environmental experiences? Or could they be more universal, perhaps resulting from mechanisms that arose over evolution and operate regardless of context or experience? We and our colleagues have begun addressing these questions by exploring whether the roots of our economic behavior – both our stable preferences and our behavioral biases – are shared by our closest living evolutionary relatives, the extant non-human primates. Since humans and capuchins are closely related biologically, yet lack similar market experience, any shared cognitive systems are likely to have a common origin.

Note, however, that our work on primate economic biases was not the first to take a principled economic approach with non-human subjects. Indeed, some elegant early work in the 1970s by Kagel and colleagues found support for the stability of preferences and the applicability of economic choice theory in standard non-human psychological subjects: rats and pigeons. In a series of studies, Kagel and colleagues trained their subjects on a lever-pressing task in which subjects had a "budget" of different lever presses, each of which delivered different rewards at different rates. The researchers then used a standard revealed-preference approach in which the subjects' choices were identified via their lever choices. Using this approach, Kagel and colleagues demonstrated that rat and pigeon behavior, like that of human consumers, appears to obey the laws of demand (Battalio *et al.*, 1981a, 1981b, 1985; Kagel *et al.*, 1975, 1981, 1990, 1995).

Unfortunately, while rats and pigeons are easy subjects to work with, their limited cognitive abilities make it difficult to investigate more subtle aspects of economic choice, including many important and systematic human biases. More importantly, rats and pigeons lack one of the hallmarks of human economies: trade. Indeed, Adam Smith famously argued that the behavior of animals was not relevant to economics because they lacked the capacity to master trade. As he put it in *The Wealth of Nations*, "Nobody ever saw a dog make a fair and deliberate exchange of one bone for another with another dog. Nobody ever saw one animal by its gestures and natural cries signify to another, this is mine, that yours; I am willing to give this for that."

Another problem with the exclusive use of rats and pigeons as models for human economic choice concerns their potential for informing claims about the evolution of human choice behavior. Although rats and pigeons are commonly used in psychological studies, they represent extremely distantly related species from an evolutionary perspective. For this reason, choice experiments involving rodents and birds are silent, both on questions regarding the evolutionary history of human choice behavior and on issues related to the neural architecture underling these behaviors. In short, although previous work with animals has adeptly demonstrated the robustness of revealed-preference techniques, the field of economics is still far from an evolutionary-history based understanding of human decision making.

The goal of our recent work on capuchin economic choice is to bridge this evolutionary divide. To do so, we have developed an experimental method for presenting choice problems to capuchin monkeys in a situation that is as analogous as possible to the markets in which humans exhibit economic choice. Before turning to these studies, we'll take a brief pause to introduce the reader to the subjects of our experiments. Since many economists (and possibly some neuroscientists) are not all that familiar with primate evolution and taxonomy, we first provide a brief introduction on the phylogenetic history of primates.

PRIMATE EVOLUTION 101

When neuroeconomists reference the brain or cognitive processes of "the monkey," they are – probably without realizing it – being incredibly imprecise. To researchers in primate cognition, the term "monkey" does not pick out a coherent natural kind – a "monkey" could mean any one of the 264 extant monkey species, all of whom inhabit different environments, eat different things, come from different genera, and presumably possess different cognitive specializations with different neural substrates (see the review in Ghazanfar and Santos, 2004). Such differences can have important consequences for the cognitive and neural capacities that these different species utilize in decision-making contexts. Even very closely related monkey species can differ drastically in fundamental cognitive processes and decision-making strategies. To take one elegant example, Stevens and colleagues (2005a, 2005b) recently observed that cotton-top tamarins (*Saguinus oedipus*) and common marmosts (*Callithrix jacchus*) – two extremely closely related New World monkey species – exhibit robust differences in their discounting behavior, with marmosets valuing future rewards more than tamarins do. As this example demonstrates, it would make little sense to talk about discounting behavior in "the monkey," as such a generalization would miss out on the fact that different kinds of monkey possess discounting functions that might be specific to their own species (or, in the case of marmosets and tamarins, specific to their species-unique feed ecology).

Typically, however, when neuroscientists refer to research with monkeys they tend to mean the species of monkey most typically used in neurophysiological studies of decision making, namely the macaque, one of several species within the genus *Macaca*[1]. Macaques are an Old World monkey species, meaning that they are native to Africa and Asia. Macaques are the mostly widely distributed genus of primates (with the exception of humans), and are thus an extremely flexible species. Because of their adaptability, macaques live well in captivity and have thus long served as a successful

---

[1] It should be remembered, however, that although macaques have predominated as neuroscientific models, some of the most important neuroscientific findings in decision making have also used a marmoset monkey model – for example, Dias *et al.*, 1996, 1997.

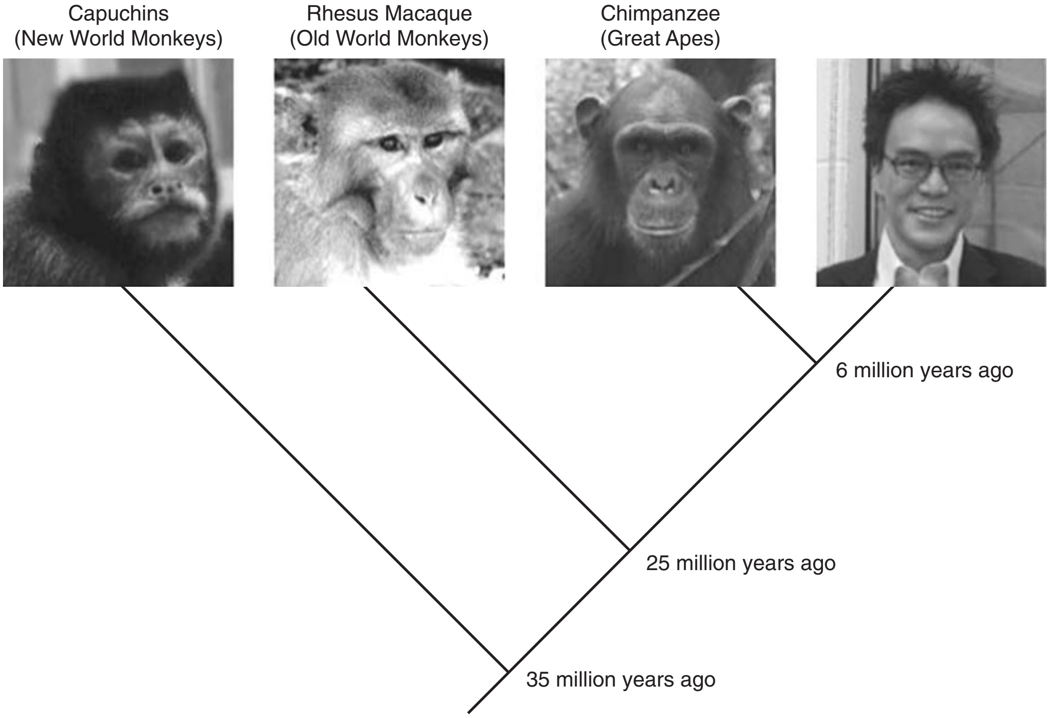


FIGURE 7.1 A schema of the primate evolutionary tree. Our subject species, the capuchin monkey, branched of from the human Old World primate line about 35 million years ago.

animal model in medical studies. Due to their prominence in early medical research, macaques were quickly imported for use in early neuroscientific investigations. Some of the first approaches to detailing the structure and function of primate motor cortex were performed on macaques in the 1800s. This early work functionally established macaques as the primate brain model for the next two centuries. Indeed, many chapters in this volume specifically focus on neuroeconomic insights gleaned from macaque brains – for example, Chapters 29 and 31.

Our behavioral work on monkey preferences does not focus on macaques, however. Instead, we work with a species believed to represent a cognitive rather than a neuroscientific model of human cognition – the brown capuchin monkey (see Chapters 18 and 19). In contrast to macaques, who are members of the Old World monkey lineage, capuchins are members of the more distantly related New World monkey branch, a group of primates that split from the Old World primate line around 35–40 million years ago (Figure 7.1). While Old World monkeys inhabit Africa and Asia, New World monkeys, like capuchins, are native to South and Central America, and thus evolved in different ecological niches than did other Old World species.

Despite millions of years of separation from our own species, the cognition of capuchin monkeys is, in many ways, quite similar to that of humans in a number of domains. Capuchins are often considered among primate researchers to be "the chimpanzee" of the New World primates. Capuchins have extremely large brains relative to their body size (see, for example, Fragaszy *et al.*, 2004a). In addition to these physical attributes, capuchins live in relatively large social groups, particularly compared to other New World species, with groups in the wild becoming as larger as 40 individuals. Despite this large group size, however, capuchins are an extremely tolerant species of primate, maintaining only a loosely defined dominance hierarchy that permits sharing food with many members of the group (de Waal, 2000; de Waal and Berger, 2000). For this reason, capuchins are extremely socially adept. Recent research suggests that they can successfully represent the goals of other individuals (Santos, personal communication) and can learn socially from the actions of others – though the specifics regarding how much they can learn continue to be debated (Adams-Curtis and Fragaszy, 1995; Custance *et al.*, 1999; Ottoni and Mannu, 2001; Visalberghi and Addessi, 2000, 2001; Brosnan and de Waal, 2004; Ottoni *et al.*, 2005; Bonnie and de Waal, 2007; see elegant reviews in Adessi and Visalberghi, 2006 and Fragaszy *et al.*, 2004b). Finally, capuchins are known for their elaborate tool-use. They use a variety of tools

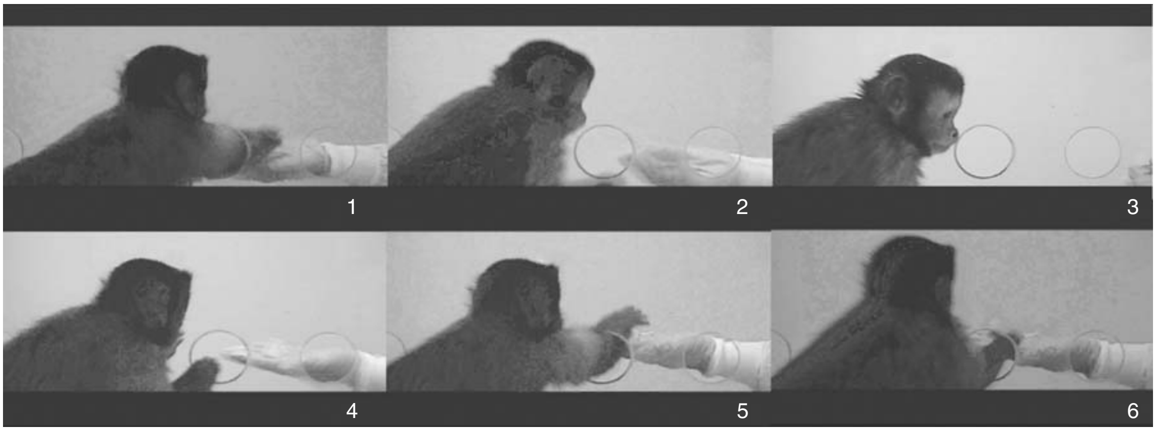


FIGURE 7.2 A frame-by-frame demonstration of a single trading event involving one of our capuchin actors (Jill). The capuchin begins by placing a token in the experimenter's hand (1). The experimenter then takes the token away (2–3) and delivers a piece of food (4) which the capuchin then takes from the experimenter's hand (5–6).

both in the wild and in captivity, including using pushing and pulling tools to gain out-of-reach food, dipping tools to gain access to out-of-reach liquids, combinations of stone hammers and anvils for opening palm nuts, and even crushed millipedes as a mosquito repellant (Fragaszy *et al.*, 2004b, Valderrama *et al.*, 2000).

REVEALING CAPUCHIN PREFERENCES: THE TOKEN TRADING METHODOLOGY

Our goal was to design a task with which we could reveal capuchins' preferences. The problem, of course, is that capuchins would presumably have some difficulty performing the tasks that experimental economists typically employ to reveal human preferences. Monkeys' preferences concerning their willingness to pay for certain gambles or bundles of goods can't be assessed using written surveys; nor can monkeys' behavior as consumers in a market be used, since they do not naturally act as consumers in markets. We therefore had to design a novel method that permitted capuchins to reveal their preferences in something like a market, a situation that was as analogous as possible to the methods used to test preferences in humans; specifically, one that involved relatively little training and also permitted formal price-theoretic analyses.

To do this, we capitalized on the fact that capuchin monkeys (as well as other primates) can be quickly trained to trade tokens for small food rewards (see, for example, Westergaard *et al.*, 1998, 2004; Liv *et al.*, 1999; Brosnan and de Waal, 2003, 2004; Adessi *et al.*, 2007). A number of different labs have successfully taught capuchins this trading methodology using an individual experimenter who would reward a capuchin subject for handing her the token. In our set-up, we hoped to give capuchins choices *between multiple different traders*, each of whom would deliver different kinds or amounts of goods when presented with a single token. In this way, we were able to put capuchins into a situation much like an economic market – one in which they could establish preferences across different bundles of goods. With this set-up, we could introduce price and wealth changes and examine how such changes affected capuchins' purchasing behavior. Further, we could observe whether capuchins preferred options that stochastically dominated all others (i.e., ones in which they unconditionally received the most food). Finally, and perhaps most importantly, we could examine whether capuchins' preferences obeyed prospect-theoretic predictions, and thus were affected by reference points and framing.

Chen *et al.* (2006) introduced five adult capuchins to this economic market. Each capuchin began testing by leaving its homecage and entering a small testing box. In the box, monkeys found a small wallet of small, disc-shaped metal tokens. Two experimenters then positioned themselves on either side of the cage. The two experimenters differed in their clothing (each wore differently colored medical scrubs) and also in the kind of good offered. On each trial, the monkey had a chance to trade a token with one of the two experimenters. Each trial began when the two experimenters were in position on either side of the cage. In one hand the experimenters held the good that they were offering to the monkey; their other hand remained open for the monkey's token (Figure 7.2). Monkeys could therefore check their options and trade with the experimenter who gave the best kind or amount of the good. Each session lasted until the monkey had spent all of its tokens.

DO CAPUCHINS OBEY PRICE THEORY AS HUMANS DO?

Our first goal was to examine whether the preferences capuchins established in the token economy we had set up mirrored those of a human economy. That is, having allocated their budget of tokens across a set of possible goods, would capuchins respond rationally to price and wealth shocks? To do this, we first found two goods that the capuchins liked equally – pieces of jello and apple slices – spending about half their budget on each of the goods. Once capuchins' choices stabilized across sessions, we introduced a compensated price shift.

In our compensated price shift, we assigned each subject a new budget of tokens and then dropped the price of one of the two goods by half. In order to respond as humans would to this price shift, capuchins must shift their consumption to the cheaper good; namely, they should spend more of their token budget on the cheaper good than they did before the price shift. The majority of our capuchins actors did just this, suggesting that they, like humans, obey the tenets of price theory.

In a further study, we examined whether capuchins also try to maximize their expected payoff in the market. If capuchins had a choice between two traders offering the same kind of good, would they choose the experimenter whose payoff stochastically dominated, the one that gave the most food overall? To look at this, we (Chen *et al.*, 2006) again presented capuchins with a choice between two traders, but this time the traders offered the same kind of good – apples. The traders differed both in the number of apples they initially offered and in the number they gave over. The first experimenter always offered the monkey one piece of apple and then handed over that one piece. The second experimenter, in contrast, was risky – he did not always hand over what he promised. This second experimenter began with two pieces of apple and then, with 50% probability, either handed over both pieces or took one of the two pieces away for an offer of only one piece. On average, however, this risky experimenter represented a good deal – he gave one-and-a-half pieces of apple on average, while the other experimenter gave only one piece. Like rational actors, our capuchin traders appeared reliably to prefer the risky experimenter who stochastically dominated. In this way, capuchins not only shift consumption rationally in response to price shifts, but also prefer trading gambles that provide the highest average payoffs.

DO CAPUCHINS DISPLAY THE SAME BIASES AS HUMANS?

Our findings that capuchins obey price theory and choose options that stochastically dominate suggest that capuchins behave rationally in their token market in some of the same ways that humans behave rationally in their economies. This work, then, set the stage for examining whether capuchins also behave *non-standardly* in the ways that humans do. Specifically, we wanted to examine whether capuchins share some of the biases that pervade human choice behavior. As decades of work in behavioral economics have shown, human consumers appear to evaluate their choices not only in terms of their expected payoffs. Instead, consumers also appear to evaluate different gambles in terms of arbitrary reference points. In particular, human participants tend to be loss averse – they avoid getting payoffs that appear as losses relative to their reference points more than they appear to seek out gains relative to their reference points (e.g., Kahneman and Tversky, 1986; Tverky and Kahneman, 1981). The phenomena of reference dependence and loss aversion have been demonstrated in countless experimental scenarios and gambles (e.g., Tversky and Kahneman, 1986), but also have demonstrated real-world manifestations in situations as diverse as unemployment patterns (Krueger and Summers, 1988; Akerlof and Yellen, 1990) housing-market changes (Odean, 1998), and asymmetric consumer elasticities (Hardie *et al.*, 1993). Further, reference dependence also affects participants' intuitions regarding fairness and moral concerns (Kahneman *et al.*, 1991).

Is reference dependence a uniquely human phenomenon, or does it extend more broadly across the animal kingdom? To examine this, we presented monkeys with trading situations in which they had the opportunity to consider their final trading payoffs relative to a reference point. We could therefore examine whether framing also affects capuchin choice and preferences.

Are Capuchins Reference Dependent and Loss Averse?

In our first study (Chen *et al.*, 2006), we explored whether capuchins, like humans, set up expectations relative to an arbitrary reference point. To do this, we independently varied what monkeys were initially shown and then what they eventually received in exchange for a token, thereby setting up situations in which the monkeys could get more or less than they expected. In the first experiment, we examined whether capuchins attended to this reference point. Monkeys got to choose between two experimenters who both delivered the same average expected payoff of one-and-a-half pieces of apples. One experimenter, however, gave this average payoff of one-and-a-half apples by way of a perceived loss. This experimenter began every trade by showing the monkeys two pieces of apple. When this experimenter was paid, he either delivered these two pieces of apple as promised, or removed one to deliver only a single apple piece. In this way, the first experimenter gave the monkeys less than they had expected based on their reference point. The second experimenter, in contrast, gave more on average than the monkeys expected. This second experimenter always began by displaying a single piece of apple but then, when paid, either delivered this one piece as promised or added a second piece for a payoff of two apple pieces. Monkeys thus had a choice of obtaining an average of one-and-a-half pieces of apple by way of a perceived loss or by way of a perceived gain. Although the average payoff was the same across the two experimenters, our monkey consumers did not prefer the two experimenters equally. Instead, they reliably preferred the experimenter who delivered his apple pieces by way of a gain. Like humans, capuchins appear to take into account reference points – in this case, what they initially are offered.

We then went on to examine whether capuchins avoid losses in the same way as humans. Did capuchins avoid the experimenter who gave them perceived losses, or did they instead seek out the experimenter who gave them perceived gains. To test this, we gave monkeys a choice between one experimenter who always delivered a loss – he consistently promised two pieces of apple and gave one – versus an experimenter who always gave what was expected – he promised one piece of apple and delivered exactly that piece. As in the previous study, our monkeys seemed to avoid the experimenter who delivered the perceived loss. Interestingly, monkeys faced with this choice robustly preferred the experimenter who gave what they expected, despite the fact that both experimenters delivered a single piece of apple on every trial.

In this way, capuchins appear to share at least two of the fundamental biases that humans display. Capuchins represent their payoffs relative to arbitrary reference points and appear to avoid gambles that are framed as losses. Such results indicate that monkeys also succumb to framing effects, with different descriptions of the same problem leading them to make different choices.

Framing and Risk: Do Capuchins Exhibit a Reflection Effect?

In our next set of studies, we examined whether framing also affects monkeys' risk preferences. To do this, we presented the capuchins with a version of Tversky and Kahneman's (1981) well-known Asian Disease problem (Lakshminarayanan, Chen, and Santos, personal communication). In each condition, monkeys had a choice between two kinds of experimenters who delivered identical expected payoffs but differed in how much their payoffs varied. Monkeys could choose to trade with a safe experimenter who traded the same way on every trial, or a risky experimenter who represented a 50–50 gamble between a high and a low payoff. What differed across the two conditions was how the experimenters framed the monkeys' choices. In the first condition, each of the experimenters framed his payoff in terms of a gain; monkeys had a choice between a safe experimenter who promised one piece of food but always delivered two, and a risky experimenter who promised one piece of food but then delivered either one piece of food or three pieces of food. Like humans tested in the Asian Disease problem, monkeys presented with gains chose to avoid risk – they reliably preferred to trade with the safe experimenter over the risky experimenter. The second condition, in contrast, presented monkeys with safe and risky losses. Monkeys had a choice between a safe experimenter who always promised three pieces of food but always delivered two, and a risky experimenter who promised three pieces of food but either delivered one piece of food or three pieces of food. In contrast to their performance in the gains condition, monkeys in the losses condition preferred to trade with the risky experimenter. In this way, monkeys appear to change their risk preferences depending on whether they are expecting perceived losses or perceived gains. Like humans, capuchins get riskier when gambling over losses than gains.

Interestingly, recent work by Kacelnik and his colleagues suggests that capuchins are not the only non-human species to show a framing-based risk-preference reversal when depending on framing; another even more distantly related non-human species – the European starling (*Sturnus vulgaris*) – shows a similar risk-preference reversal on an analogous choice task. Marsh and Kacelnik (2002) presented starlings with a task in which they could choose either fixed or variable rewards. Starlings practiced this task with one expected payoff amount, and were then tested with outcomes that were either more or less than their expectations. Starlings preferred the risky option more when they received less than they expected rather than when they received more than they expected, suggesting that starlings also become more risk-prone when dealing with perceived losses than with perceived gains. Combined with our capuchin studies, this work suggests that framing effects may extend broadly across the animal kingdom, and may also extend to a variety of taxa.

Do Capuchins Exhibit An Endowment Effect?

We then went on to examine whether capuchins demonstrate an endowment effect (see Thaler, 1980), a phenomenon in which ownership increases an object's value. In what is now a classic paper, Kahneman *et al*. (1990) presented half of a group of human participants with a coffee mug, and then allowed participants to either buy or sell the new mug. Kahneman and colleagues found that participants that owned the mug demanded a higher price to sell the mug than non-owners would pay to buy it. This discrepancy between owners' willingness-to-accept and buyers' willingness-to-pay was christened the *endowment effect*. Although there is still considerable debate concerning the exact mechanisms underlying the endowment effect, many have hypothesized that this effect follows from loss aversion (see Kahneman *et al*., 1990). If this is the case, then capuchins – who exhibit loss aversion in our experimental market – may also show a bias towards over-valuing objects that they own over those they don't yet own.

In a recent study (Lakshminarayanan, Chen, and Santos, personal communication), we explored whether capuchins were also susceptible to endowment effects (see Chapter 19 for similar experiments with chimpanzees). We first determined two goods that the monkeys preferred about equally, splitting their budget of tokens across the two goods. We then made our capuchin subjects the "owners" of one of the two equally preferred goods. Rather than giving each monkey subject a wallet of tokens, we instead provided a wallet of goods and allowed them to trade for the other equally preferred good. Since the two goods were already shown to be equally preferred, it might be predicted that capuchins would trade about half their endowed goods and then keep the other half. However, in contrast to this prediction, our capuchin actors reliably preferred to keep the food with which they were endowed. Control conditions revealed that our observed effect was not due timing effects or transaction costs – monkeys failed to trade their endowed good even in cases in which they were compensated for the cost of the trade and the time it takes to wait for the trade to be completed. These results provide the first evidence to date that a non-human species demonstrates a true endowment effect – one that cannot be explained by timing, inhibition, or problems with transaction-related costs.

WHAT COMPARATIVE WORK MEANS FOR TRADITIONAL ECONOMICS AND NEUROECONOMICS

When taken together, our comparative studies to date suggest that capuchin monkey preferences operate in much the same way as those of human agents. First, capuchins appear to obey the standard tenets of price theory, just like humans. In spite of their obedience to price theory, however, capuchins also exhibit the same systematic biases as humans – they evaluate gambles in terms of arbitrary reference points, and pay more attention to losses than to gains. Finally, monkeys appear to show other market anomalies, like the endowment effect. Our work thus suggests that human behavioral biases result not from species-unique market experiences or cultural learning; instead, such biases are more likely to be far more basic, perhaps even evolved strategies present long ago in our common ancestor with other New World primate species.

This work further suggests that such biases may emerge in the absence of much market experience not just in capuchins, but in the human species as well. Indeed, our work provides hints about another possible and probably fruitful line of work on the origins of preference. Our studies to date have focused on the evolutionary origins of human preferences and incentives, but even less work has examined the *ontogenetic origins* of these phenomena – namely, how they develop over the human lifecourse (for review, see Santos and Lakshminarayanan, 2008). Although some work to date has examined the development of loss aversion (e.g., Reyna and Ellis, 1994) and the endowment effect (see Harbaugh *et al*., 2001) in children, there is still relatively little consensus concerning how and when behavioral biases emerge in human decision making. In addition, to our knowledge, all of the available evidence to date examining the development of revealed preferences has involved older children – participants who've had at least some experience making purchases in the real world. For this reason, older children are not the best subject pool if intending to examine the role of experience in the development of loss aversion and reference dependence. To really study the role of experience, researchers should focus their empirical effort on studying human infants – participants who are young enough to lack *any* market experience. Although human infants' preferences are not currently a standard focus for economic experimentation, there is no reason they cannot become one. In the past decade, developmental psychologists have established a number of empirical methods that can be easily imported for use in economic studies with preverbal infants. Infant-researchers have developed standard methods for assessing both infants' choices (e.g., Feigenson *et al.*, 2002) and their preferences (e.g., Spelke, 1976), all using non-verbal techniques. Using these experimental methods, economists could potentially ask whether infants obey price theory (and thus examine whether an obedience to price theory can emerge in the complete absence of experience). Similarly, it would be possible to examine how and when biases like loss aversion and reference dependence begin emerging, and again explore the role of economic experience and other factors in the development of these heuristics.

Our finding that many behavioral biases are shared by closely related primates has a number of implications for practicing economists. The first of these involves how an economist might choose to treat behavioral biases in both positive and normative terms. For example, if biases observed in human behavior are the results of misapplied heuristics, then it seems natural to assume that what is learned can be unlearned, and that these mistakes are likely to disappear quickly in the face of market pressures – especially when stakes are high. Our work, however, suggests that these biases emerge in the absence of experience, and thus that biases are likely to manifest themselves in novel situations. Such findings also raise the hurdle that competitive pressure may need to pass to discipline behavior. From a positivist perspective, while it may still be reasonable to believe that in high-stakes settings where market participants are exposed to constant feedback markets will display extremely rational behavior, those settings might not represent the majority of economically relevant settings. Indeed, consistent with classical welfare analysis, if a bias repeatedly emerges in different market settings and represents a fundamental aspect of people's preferences, it may demand more normative weight than we might have otherwise thought.

Our work also has important implications for non-traditional economists – neuroeconomists interested in the neural basis of standard and non-standard economic behavior. In the past decade, macaque models have afforded neurophysiologists with a number of important discoveries concerning the neural basis of our representation of risk and value (see Chapters 29, 31, and 32). Many of the neurophysiological studies to date, however, have concerned themselves with more standard aspects of choice behavior. In contrast, fMRI research with humans has focused on the neural basis of more non-standard behaviors, namely behavioral biases. While such fMRI techniques have already provided tremendous insight into the neural basis of these framing effects (see, for example, Chapters 10 and 11), these methods would undoubtedly be complemented by neurophysiology work examining framing effects at the level of individual neurons. To date, however, little neurophysiological work has addressed the role of context and framing, in part because designing framing tasks for use in non-verbal primate subjects is a non-trivial task. The trading experiments we have developed for capuchins, however, demonstrate that such framing effects can and do occur in a non-verbal species. Our work suggests that a physiological investigation of framing is possible, and thus that it might be possible to examine prospect theoretic predictions in a primate neural model. Our work demonstrating that monkeys exhibit an endowment effect further suggests that physiologists might be able to examine even more subject changes in valuation – such as those due to ownership – in a primate model as well.

The field of neuroeconomics, though still in its infancy, has enjoyed much success in a relatively short amount of time. Undoubtedly, much of the success of this newly emerging field relies on the importance it places on interdisciplinary approaches to the study of economic behavior. Our goal in this chapter has been to point out how primate cognition studies of choice, preferences, and incentives can add to this empirical mix – both in their own right as a way of examining the origins of standard and non-standard economic behavior, and for their potential to give rise to new behavioral assays needed for neurophysiological insights into human economic behavior.

Acknowledgements

The authors would like to thank Venkat Lakshminarayanan, members of the Yale Neuroeconomics Group, and the editors for helpful comments on this chapter. We are also grateful to Brian Hare and Vanessa Woods for use of their chimpanzee photo. This work was supported by a National Science HSD Award (#0624190), the Whitebox Advisors, the Russell Sage Foundation, and Yale University. Address correspondence to laurie.santos@yale.edu.

References

Adams-Curtis, L. and Fragaszy, D.M. (1995). Influence of a skilled model on the behavior of conspecific observers in tufted capuchin monkeys (*Cebus apella*). *Am. J. Primatol.* 37, 65–71.

Addessi, E. and Visalberghi, E. (2006). Rationality in capuchin monkeys feeding behaviour? In: S. Hurley and M. Nudds (eds), *Rational Animals?* Oxford: Oxford University Press, pp. 313–328.

Addessi, E., Crescimbene, L., and Visalberghi, E. (2007). Do capuchin monkeys (*Cebus apella*) use tokens as symbols? *Proc. R. Soc. Lond. Series B* 274, 2579–2585.

Akerlof, G.A. and Yelllen, J.L. (1990). The fair wage–effort hypothesis and unemployment. *Q. J. Economics* 105, 255–283.

Battalio, R.C., Kagel, J.H., Rachlin, H., and Green, L. (1981a). Commodity-choice behavior with pigeons as subjects. *J. Political Econ.* 89, 67–91.

Battalio, R.C., Green, L., and Kagel, J.H. (1981b). Income–leisure tradeoffs of animal workers. *Am. Econ. Rev.* 71, 621–632.

Battalio, R.C., Kagel, J.H., and MacDonald, D.N. (1985). Animals' choices over uncertain outcomes: some initial experimental results. *Am. Econ. Rev.* 75, 597–613.

Becker, G.S. (1962). Irrational behavior and economic theory. *J. Political Econ.* 70, 1–13.

Becker, G.S. (1976). *The Economic Approach to Human Behavior*. Chicago, IL: University of Chicago Press.

Becker, G.S. and Murphy, K.M. (1988). A theory of rational addiction. *J. Political Econ.* 96, 675–700.

Bonnie, K.E. and de Waal, F.B.M. (2007). Copying without rewards: socially influenced foraging decisions among brown capuchin monkeys. *Animal Cogn.* 10, 283–292.

Brosnan, S.F. and de Waal, F.B.M. (2003). Monkeys reject unequal pay. *Nature* 425, 297–299.

Brosnan, S.F. and de Waal, F.B.M. (2004). Socially learned preferences for differentially rewarded tokens in the brown capuchin monkey (*Cebus apella*). *J. Comp. Psychol.* 118, 133–139.

Camerer, C.F. (1998). Bounded rationality in individual decision making. *Exp. Economics* 1, 163–183.

Chen, M.K., Lakshminaryanan, V., and Santos, L.R. (2006). The evolution of our preferences: evidence from capuchin monkey trading behavior. *J. Political Econ.* 114, 517–537.

Custance, D.M., Whiten, A., and Fredman, T. (1999). Social learning of "artificial fruit" processing in capuchin monkeys (*Cebus apella*). *J. Comp. Psychol.* 113, 13–23.

de Waal, F.B.M. (2000). Primates – a natural heritage of conflict resolution. *Science* 289, 586–590.

de Waal, F.B.M. and Berger, M.L. (2000). Payment for labour in monkeys. *Nature* 404, 563.

Dias, R., Robbins, T.W., and Roberts, A.C. (1996). Dissociation in prefrontal cortex of affective and attentional shifting. *Nature* 380, 69–72.

Dias, R., Robbins, T.W., and Roberts, A.C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analogue of the Wisconsin Card Sort Test: restriction to novel situations and independence from "on-line" processing. *J. Neurosci.* 17, 9285–9297.

Feigenson, L., Carey, S., and Hauser, M. (2002). The representations underlying infants' choice of more: object files versus analog magnitudes. *Psychological Sci.* 13, 150–156.

Fragaszy, D., Visalberghi, E., and Fedigan, L. (2004a). *The Complete Capuchin*. Cambridge: Cambridge University Press.

Fragaszy, D.M., Izar, P., Visalberghi, E., Ottoni, E.B., and Gomes De Oliveira, M. (2004b). Wild capuchin monkeys (*Cebus libidinosus*) use anvils and stone pounding tools. *Am. J. Primatol.* 64, 359–366.

Ghazanfar, A.A. and Santos, L.R. (2004). Primate brains in the wild: the sensory bases for social interactions. *Nat. Rev. Neurosci.* 5, 603–616.

Glaeser, E.L. (2004). Psychology and the market. *Am. Econ. Rev.* 94, 408–413.

Gul, G. and Pesendorfer, W. (2001). Temptation and self-control. *Econometrica* 69, 1403–1435.

Harbaugh, W.T., Krause, K., and Vesterlund, L. (2001). Are adults better behaved than children? Age, experience, and the endowment effect. *Econ. Letts* 70, 175–181.

Hardie, B.G.S., Johnson, E.J., and Fader, P.S. (1993). Modeling loss aversion and reference dependence effects on brand choice. *Marketing Sci.* 12, 378–394.

Kagel, J.H., MacDonald, D.N., Battalio, R.C. *et al*. (1975). Experimental studies of consumer demand behavior using laboratory animals. *Economic Inquiry* 13, 22–38.

Kagel, J.H., Battalio, R.C., and Rachlin, H. (1981). Demand curves for animal consumers. *Q. J. Economics* 96, 1–16.

Kagel, J.H., MacDonald, D.M., and Battalio, R.C. (1990). Tests of "fanning out" of indifference curves: results from animal and human experiments. *Am. Econ. Rev.* 80, 912–921.

Kagel, J.H., Battalio, R.C., and Green, L. (1995). *Economic Choice Theory: An Experimental Analysis of Animal Behavior*. Cambridge: Cambridge University Press.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–292.

Kahneman, D., Slovic, P., and Tversky, A. (eds) (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Kahneman, D., Knetsch, J.L., and Thaler, R.H. (1990). Experimental tests of the endowment effect and the Coase theorem. *J. Political Econ.* 98, 1325–1348.

Kahneman, D., Knetsch, J.L., and Thaler, R.H. (1991). Anomalies: the endowment effect, loss aversion, and status quo bias. *J. Econ. Persp.* 5, 193–206.

Kreps, D.M. and Porteus, E.L. (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica* 46, 185–200.

Krueger, A.B. and Summers, L.H. (1988). Efficiency wages and the inter-industry wage structure. *Econometrica* 56, 259–293.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Q. J. Economics* 112, 443–477.

Liv, C., Westergaard, G.C., and Suomi, S.J. (1999). Exchange and value in *Cebus apella*. *Am. J. Primatol.* 49, 74–75.

Marsh, B. and Kacelnik, A. (2002). Framing effects and risky decisions in starlings. *Proc. Natl Acad. Sci.* 99, 3352–3355.

Odean, T. (1998). Are investors reluctant to realize their losses? *J. Finance* 5, 1775–1798.

Ottoni, E.B. and Mannu, M. (2001). Semi-free ranging tufted capuchins (*Cebus apella*) spontaneously use tools to crack open nuts. *Intl J. Primatol.* 22, 347–357.

Ottoni, E.B., de Resende, B.D., and Izar, P. (2005). Watching the best nutcrackers: what capuchin monkeys (*Cebus apella*) know about others' tool-using skills. *Animal Cogn.* 8, 215–219.

Reyna, V.F. and Ellis, S.C. (1994). Fuzzy-trace theory and framing effects in children's risky decision making. *Psychological Sci.* 5, 275–279.

Santos, L.R. and Lakshminarayanan, (2008). Innate constraints on judgment and choice? Insights from children and non-human primates. In: P. Carruthers (ed.), *The Innate Mind: Foundations and The Future*. Oxford: Oxford University Press, in press.

Spelke, E.S. (1976). Infants' intermodal perception of events. *Cogn. Psychol.* 8, 553–560.

Stevens, J.R., Hallinan, E.V., and Hauser, M.D. (2005a). The ecology and evolution of patience in two New World primates. *Biol. Letts* 1, 223–226.

Stevens, J.R., Rosati, A.G., Ross, K.R., and Hauser, M.D. (2005b). Will travel for food: spatial discounting and reward magnitude in two New World monkeys. *Curr. Biol.* 15, 1855–1860.

Thaler, R.H. (1980). Toward a positive theory of consumer choice. *J. Econ. Behav. Org.* 1, 39–60.

Tversky, A. and Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science* 211, 453–458.

Tversky, A. and Kahneman, D. (1986). Rational choice and the framing of decisions. *J. Business* 59, 251–278.

Valderrama, X., Robinson, J.G., Attygalle, A.B., and Eisner, T. (2000). Seasonal anointment with millipedes in a wild primate: a chemical defense against insects? *J. Chem. Ecol.* 26, 2781–2790.

Visalberghi, E. and Addessi, E. (2000). Seeing group members eating a familiar food affects the acceptance of novel foods in capuchin monkeys, *Cebus apella*. *Animal Behav.* 60, 69–76.

Visalberghi, E. and Addessi, E. (2001). Acceptance of novel foods in capuchin monkeys: do specific social facilitation and visual stimulus enhancement play a role? *Animal Behav.* 62, 567–576.

Westergaard, G.C., Liv, C., Chavanne, T.J., and Suomi, S.J. (1998). Token mediated tool-use by a tufted capuchin monkey (*Cebus apella*). *Animal Cogn.* 1, 101–106.

Westergaard, G.C., Liv, C., Rocca, A. *et al*. (2004). Capuchin monkeys (*Cebus apella*) attribute value to foods and tools during voluntary exchanges with humans. *Animal Cogn.* 7, 19–24.

CHAPTER

8

The Expected Utility of Movement

Julia Trommershäuser, Laurence T. Maloney and Michael S. Landy

OUTLINE

INTRODUCTION

Our survival depends on our ability to act effectively, maximizing the chances of achieving our movement goals. In the course of a day we make many movements, each of which can be carried out in a variety of ways. Shall I reach for that wine glass quickly or slowly? Shall I approach from the right or left? Movement planning is a form of decision making as we choose one of many possible movement strategies to accomplish any given movement goal. It is important for us to make these "motor decisions" rapidly and well. In this chapter, we consider how movements are planned and show that a certain class of movement-planning problems is mathematically equivalent to a choice among lotteries in decision making under risk or ambiguity (see also Trommershäuser *et al.*, 2006a). This analogy allows us to examine movement planning from a new perspective, that of the ideal economic movement planner. It also allows us to contrast how we make decisions in two very different modalities; planning of movement, and traditional economic decision making.

We review our previous work on movement-planning under risk, in which subjects are generally found to be very good at choosing motor strategies that come close to maximizing expected gain – a result that is in contrast with that found with paper-and-pencil decision-making tasks. We discuss the implications of these different behavioral outcomes, noting the evident differences between the sources of uncertainty and how information about uncertainty is acquired in motor and economic tasks. We finally review the evidence concerning the neural coding of probability, expected movement error, and expected gain in movement under risk (see also Chapters 23, 30, and 32). We begin, however, with a brief review of previous work on how biological organisms plan movement.

MOVEMENT PLANNING AS OPTIMIZATION

In planning a movement, the brain has to select one of many possible movement plans or strategies. The result of executing a movement strategy is an actual trajectory, and it is evidently desirable that the choice of strategy satisfies the demands of the particular task and also minimizes "wear and tear" on the organism.

Typical research in the field of human motor control combines theoretical and experimental approaches. For example, a participant in a motor control experiment might perform simple reaching movements to a target, often constrained to two dimensions – for example, along the top of a table (Figure 8.1a). The recorded movement trajectories are compared with the predictions of a computational model mimicking the conditions of the experiment. Early approaches to modeling movement planning take the form of an optimization problem in which the cost function to be minimized is biomechanical and the optimization goal is to minimize some measure of stress on the muscles and joints. These models differ primarily in the choice of the cost function.

Movement Planning Optimizing Biomechanical Constraints

Possible biomechanical cost functions include measures of joint mobility (Soechting and Lacquaniti, 1981; Kaminsky and Gentile, 1986), muscle tension changes (Dornay *et al.*, 1996), mean squared rate of change of acceleration (Flash and Hogan, 1985), mean torque change (Uno *et al.*, 1989), total energy expenditure (Alexander, 1997), and peak work (Soechting *et al.*, 1995). The outcome of applying these models is typically a single, deterministic trajectory that optimizes the tradeoff between the goal of the movement and the biomechanical costs for the organism. These models are successful in explaining the human ability to adapt to forces applied during movement execution (Burdet *et al.*, 2001; Franklin *et al.*, 2007). Although this wide variety of cost functions has been employed, nearly all have successfully modeled reaching movements as following a nearly straight path with a bell-shaped velocity profile (Figure 8.1). A combination of biomechanical constraints was demonstrated by Cuijpers *et al.* (2004), who showed that humans will grasp an elliptical cylinder along either its major or its minor axis, resulting in a stable grip, but will tend to choose the axis closer to that used for circular cylinders (i.e., the grasp that is more comfortable).

Compensation for Noise and Uncertainty

The models just described yield single deterministic trajectories that do not take into account the possible consequences of motor errors due to noise in the motor system. Consider two movement plans for a reach toward a target. The first involves maximal acceleration toward the target for half the reach, and maximal deceleration to come to a halt at the target. The second has a smooth acceleration/deceleration profile. We could consider the biomechanical costs associated with the two movements, but we can also characterize the two movements in terms of an external criterion of success: which is more likely to hit the target? In moving from the purely internal criterion of success (minimizing biomechanical costs) to an external measure, we change the nature of the movement-planning problem and its solution.

Figure 8.2 shows two movement plans, in this case for a 10° deg saccadic eye movement, both of which take the same amount of time to arrive, on average, at the identical target location (Harris and Wolpert, 1998). The two planned movements differ in how force is distributed along the reach. Movement 1 (Figure 8.2, dashed curves) begins with an explosive burst of force, rapidly rotating the eye toward the target, and ends with an equal but opposite explosive burst as the eye approaches the target. Such an extreme use of force is often referred to as "bang–bang" control. Movement 2 (solid curves) has a smooth acceleration profile, with gradual acceleration during the first half of the reach and deceleration during the second.

Uncertainty in the motor system originates from noisy neural control signals that lead to variable motor output. The noise is signal-dependent; larger control signals lead to larger variability in motor outcome (Harris and Wolpert, 1998). Thus, faster movements, which require larger control signals, are more variable, resulting in the well-known speed–accuracy tradeoff. Modeling such signal-dependent noise, Harris and Wolpert (1998) found that the movement that

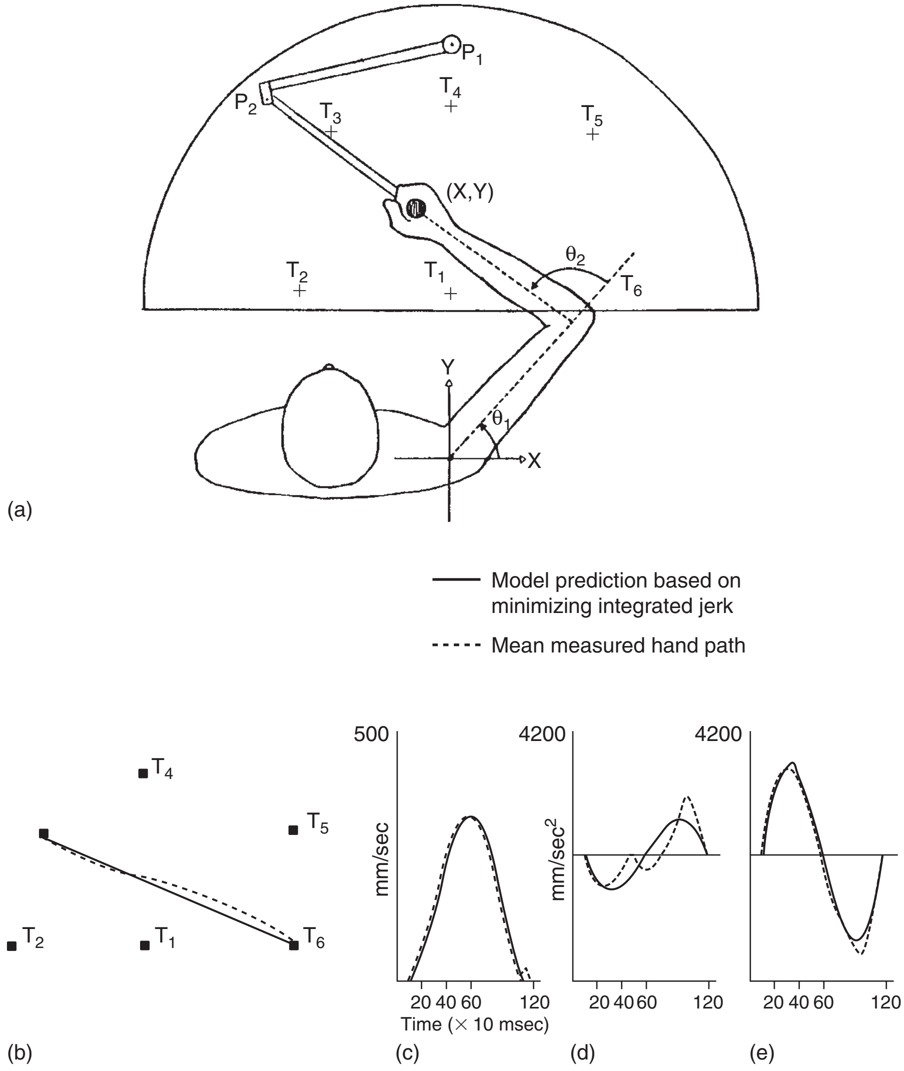


FIGURE 8.1 Example of a model of movement planning as optimization of a biomechanical constraint. (a) Subjects moved a manipulandum between various target positions (labeled T1, …, T6). (b) Hand path toward one example target. (c) Hand velocity. (d–e) Acceleration in the *x* (panel (d)) and *y* (panel (e)) directions. Dashed curves, mean measured hand path; solid curves, hand paths predicted by a model that minimized integrated jerk (rate of change of acceleration). The measured hand path shows a typical smoothly curved trajectory with a bell-shaped velocity profile. The minimum-jerk model (dashed curves) does a reasonable job of predicting movement kinematics. Reprinted from Flash and Hogan (1985), with permission.

minimizes positional variance at the end of the movement, subject to the two constraints of movement duration and mean endpoint, is Movement 2 (solid curves).

Recent experimental work concerned with the planning and execution of speeded eye and arm movements indicates that the complex sequences of neural events that underlie voluntary movements are selected so as to minimize movement error (Sabes and Jordan, 1997; Harris and Wolpert, 1998; Todorov, 2004). Note that this approach is based on the notion that the endpoint variability is a consequence of the "biological noise" in the motor control system, and therefore unavoidable.

Following the observation that movements are corrupted by considerable motor noise and do not always follow the same deterministic trajectory, Harris and Wolpert (1998) suggested that movement trajectories are selected to minimize the variance of the final eye or arm position. They proposed that the underlying determinant of trajectory planning is the minimization of the noise in the neural control signal that activates the muscles during the execution of a motor command and in the post-movement period. In their model, the final arm or eye position is computed as a function of a (deterministic) biomechanical expression and a noisy neural signal, where the noise increases with the magnitude of the neural signal (Figure 8.2). According to the model of Harris and Wolpert (1998), the planned trajectory of the eye and arm is chosen to minimize variance in the endpoint of the movement. The idea behind this approach is that the variability in the final position of a saccade or pointing movement is the result of the accumulated deviations of the executed trajectory from the planned trajectory over the duration of the movement. Horizontal saccadic eye movements, hand paths for a set of point-to-point

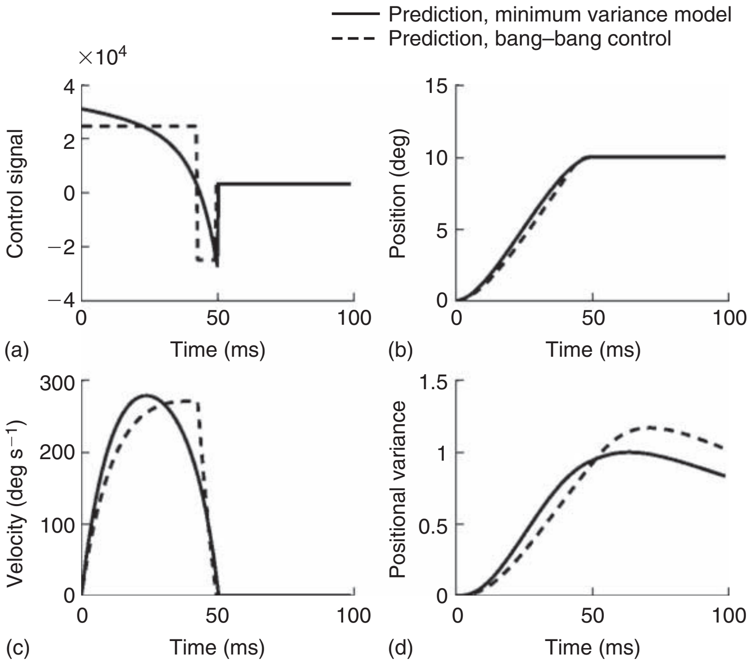


FIGURE 8.2 Predictions of a model of a 10°, 50-ms eye movement with signal-dependent noise. Dashed curves, the result of "bang–bang" control in which nearly maximum acceleration is applied throughout the movement until just prior to movement completion, at which point the eye is quickly brought to a stop; solid curves, movement achieving minimum positional variance at the end of the movement. The minimum-variance model predicts a bell-shaped velocity profile and results in substantially lower end-point variance. (a) Control signal; (b) position; (c) velocity; (d) positional variance.
Reprinted from Harris and Wolpert (1998), with permission.

movements (Harris and Wolpert, 1998), and the movement trajectories measured in an obstacle-avoidance task (Hamilton and Wolpert, 2002) are all consistent with predictions of the model.

In a similar approach, Sabes and Jordan (1997) studied the contribution of kinematic and dynamic properties of the arm to movement planning. Subjects moved their hands between pairs of targets, avoiding an obstacle along the path. To carry out this task rapidly while avoiding the obstacle, subjects should choose a trajectory so that the direction from the obstacle to the trajectory at its nearest approach is one for which movement variability is minimal. Given the difficulty of measuring two-dimensional movement variability during a reach, the authors instead modeled the sensitivity of the arm in various directions and at various positions in the workspace. They defined three sensitivity measures (kinematic, inertial, and elastic), each of which provided predictions of the point of closest approach to the obstacle that was most safe. The data were in qualitative agreement with the predictions of all three sensitivity measures, and best predicted by their inertial sensitivity model (*mobility*).

Optimization of the Consequences of Movement

So far, we have considered the costs of movement in terms of biomechanical costs (energy, wear and tear) and movement accuracy. Both these criteria for movement planning may be used by humans for planning movements. We propose that these costs may be additive and that each may be "priced" in terms of utility for the movement planner. However, in most movements there will be utilities that are a function of the movement outcome itself. A reach for a wine glass that succeeds and leads to a more rapid drink may be pleasant, but the consequences of failure (wine spilled on the new carpet or a broken glass owing to a collision along the way) may offset any desire to speed up the reach. Thus, we seek a way to frame the tradeoff between a small probability of a collision with an obstacle and a large decrease in the chances of achieving the goal of the task. Why might a subject be more willing to risk collision with an obstacle as the reward associated with successful completion of the task is increased?

There is considerable evidence that the motor system takes its own uncertainty into account when planning movements. Consider the task of moving the hand quickly to a target. The task is more difficult for shorter movement times, and for smaller and more distant targets. Subjects consistently prolonged their movement time for smaller target diameters (Fitts, 1954). Under natural reaching conditions, subjects take the uncertainty associated with their movement into account and select movement times that allow the target to be hit with constant reliability (Meyer *et al.*, 1988).

In our own work on movement-planning under risk, we make the connection between movement outcome and its consequences explicit. In our movement tasks, subjects receive monetary rewards based on the

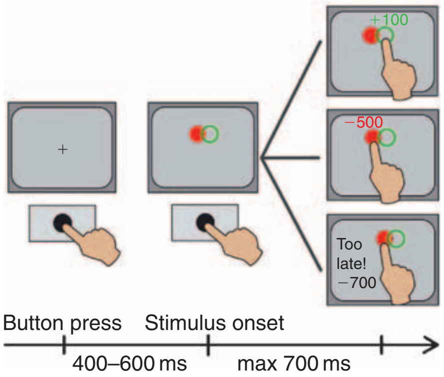


FIGURE 8.3 A movement task equivalent to decision making under risk. Subjects were required to touch a computer screen within a short period of time (e.g. 700 ms). Subjects won 100 points by hitting inside the green circle and lost 500 points by hitting inside the red circle. Subjects did not win or lose points by hitting the background, as long as the movement was completed before the time limit, but a large penalty (700 points) was imposed for movements that were too slow. Because movements were rapid, they were also variable: the finger did not always land where the subject intended. As a result, it was in the best interest of the subject to aim to the right of the center of the green circle, compromising between the loss of points due to occasionally missing the green target circle and the loss of points due to occasionally landing in the red penalty circle.

outcome of their hand movement. The central difficulty for the movement planner in these tasks is that, with speeded movement, planned movement will differ from actual movement due to motor noise. As a result, the exact outcome of the hand movement is stochastic and the choice of a movement plan simply serves to assign probabilities to each movement outcome.

In our studies, subjects pointed rapidly at stimulus configurations consisting of a small target and nearby penalty region (Figure 8.3). Hand movements that ended within the green target circle yielded a small monetary reward; those ending in the red penalty circle could result in a loss. Endpoints in the target–penalty overlap region led to the awarding of both the reward and the penalty. A time limit was imposed, and movements that arrived after the time limit resulted in large penalties. Target size and the distance between target and penalty regions were small (<2 cm), similar in size to the subject's movement endpoint variability. The movement plan that maximizes expected gain under these conditions depends on the relative position of the target and penalty circle, on the loss assigned to the penalty region, and on the subject's endpoint variability; this is explained next.

How should a subject perform this task? Clearly, the subject's visuo-motor strategy should take into account motor uncertainty and the penalty structure imposed by the task. Our model of optimal performance is built on the following assumptions:

1. When the motor system selects a visuo-motor strategy, it in effect imposes a probability density on the space of possible movement trajectories that could occur once the motor strategy is executed. This probability density is likely affected by the goal of the movement, the planned duration, the possibility of visual feedback during the movement, previous training, and intrinsic uncertainty in the motor system (see, for example, Tassinari *et al.*, 2006; Dean *et al.*, 2007). We emphasize that the consequences for the subject are completely mediated through this probability density, and we can, for the most part, ignore the details of the actual mechanisms that produce and steer the action.
2. Whatever the penalty structure of the task, the penalty incurred by the subject depends only on the motion trajectory that *actually* occurs.
3. The subject acts so as to produce maximum expected gain (MEG) as computed from the magnitude of each possible reward and penalty and the probability of incurring it.

According to this model, the goal of movement planning is to select an optimal visuo-motor movement strategy (i.e. a movement plan) that specifies a desired movement trajectory. In this model, the optimal movement strategy is the one that maximizes expected gain. The model takes into account explicit gains associated with the possible outcomes of the movement, the mover's own task-relevant variability, and costs associated with the time limits imposed on the mover.

For the conditions of this experiment, the scene is divided into four regions (Figure 8.4a): the reward-only region R_1 with gain G_1,[1] the overlap region R_2 with gain G_2, the penalty-only region R_3 with gain G_3, and the background region R_4 with gain $G_4 = 0$. We define an optimal visuo-motor strategy S as one that maximizes the subject's expected gain

$$\Gamma(S) = \sum_{i=1}^{4} G_i \, P(R_i \mid S) + G_{\text{timeout}} P(\text{timeout} \mid S) . \tag{8.1}$$

Here, $P(R_i \mid S)$ is the probability, given a particular choice of strategy S, of reaching region R_i before the time limit ($t = \text{timeout}$) has expired,

---

[1] Here we refer to outcomes as gains denoted G_i with losses coded as negative gains. The term "expected gain" that we use corresponds exactly to expected value in the psychological and economic literature.

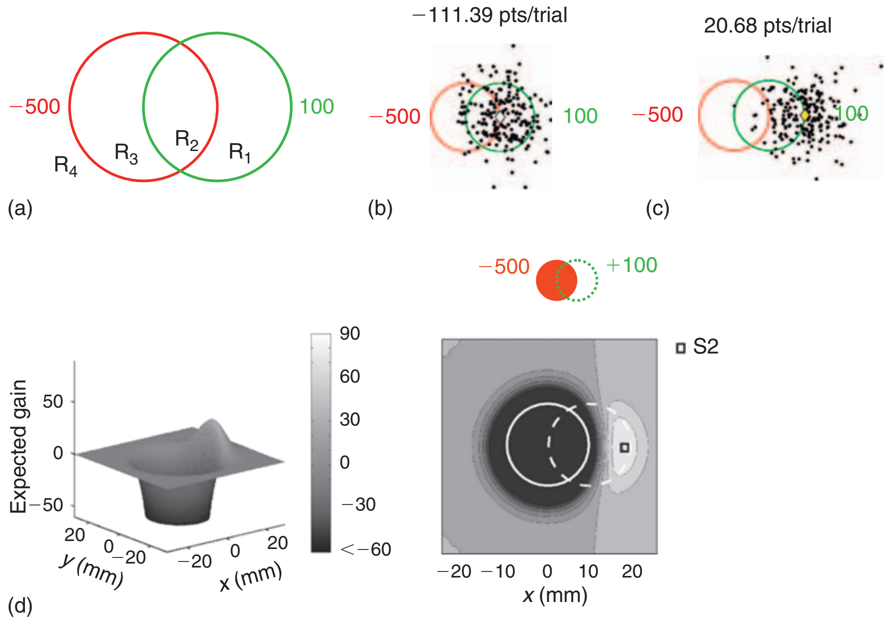


FIGURE 8.4 Equivalence between movement under risk and decision making under risk. (a) There are four regions in which the endpoint can land in the task outlined in Figure 8.3: reward-only (region with expected gain +100), reward–penalty overlap (region with expected gain −400), penalty-only (region with expected gain −500) and background (region with expected gain 0). (b) Sample endpoints for a subject aiming at the center of the green target (aim point indicated by the white diamond). This subject had a motor uncertainty of 5.6 mm (standard deviation); target and penalty circles have radii of 9 mm. The combination of motor uncertainty and aim point specifies the probability of landing in each region. This movement strategy yields an expected gain of −111.39 points/trial due to the high frequency of touching inside the red penalty circle. (c) Sample endpoints for the same subject aiming to the right of the target center to avoid accidental hits inside the penalty circle. The expected gain = 20.68 points/trial corresponds to the optimal strategy maximizing expected gain. (d) Expected gain for this subject as a function of mean movement endpoint. The maximum of this function corresponds to the aim point illustrated in (c).

$$P(R_i|S) = \int_{R_i^{\text{timeout}}} P(\tau|S)d\tau, \qquad (8.2)$$

where R_i^{timeout} is the set of trajectories τ that pass through R_i at some time after the start of the execution of the visuo-motor strategy and before the timeout. The task involves a penalty for not responding before the time limit (G_{timeout}). The probability that a visuo-motor strategy S leads to a timeout is $P(\text{timeout}\,|\,S)$.

In our experiments, subjects win or lose points by touching the reward and penalty regions on the plane of the display before the timeout. Penalties and rewards depend only on the position of the endpoint in this plane, so a strategy S can be identified with the mean endpoint on the plane $(\bar{x},\bar{y})$ that results from adopting a particular choice of strategy S that results in that mean endpoint.

In most of our experiments, subjects' movement variance was the same in the vertical and horizontal directions, indistinguishable from a bivariate Gaussian distribution (see Figure 8.4b, 8.4c for simulated data for two aim points), and remained stable throughout the experiment (see, for example, Trommershäuser *et al.*, 2003a, 2005; Wu *et al.*, 2006). Thus, we assume that the movement endpoints (x, y) are distributed according to a spatially isotropic Gaussian distribution with standard deviation σ,

$$p(x,y\,|\,\bar{x},\bar{y},\sigma^2) = \frac{1}{2\pi\sigma^2}\exp(-((x-\bar{x})^2 + (y-\bar{y})^2)/2\sigma^2). \qquad (8.3)$$

The probability of hitting region R_i is then

$$P(R_i|\bar{x},\bar{y},\sigma^2) = \int_{R_i} p(x,y\,|\,\bar{x},\bar{y},\sigma^2)dxdy. \qquad (8.4)$$

In our experiments, the probability of a timeout is effectively constant over the limited range of relevant screen locations so, for a given endpoint variance σ^2, finding an optimal movement strategy corresponds to choosing a strategy with mean aim point $(\bar{x},\bar{y})$ that maximizes

$$\Gamma(\bar{x},\bar{y}) = \sum_{i=1}^{4} G_i P(R_i | \bar{x},\bar{y},\sigma^2). \tag{8.5}$$

The maximum of $\Gamma(\bar{x},\bar{y})$ corresponds to the strategy maximizing expected gain, and depends on the position and magnitude of the penalty and on the distribution of the subject's endpoints (Figure. 8.4d). When the penalty is zero, the aim point maximizing expected gain (and hence the mean endpoint maximizing expected gain) is the center of the target region. When the penalty is non-zero, the aim point maximizing expected gain shifts away from the penalty region and, therefore, away from the center of the target. This optimal shift is larger for greater penalties, for penalty regions closer to the target, and for larger magnitudes of motor variability.

For all conditions, we compared subjects' mean endpoints with those of a movement planner that maximized expected gain by taking into account its own task-relevant variability. Once we had measured the task-relevant variability for each subject and for each condition, our model yielded parameter-free predictions of behavior maximizing expected gain for all experimental conditions against which subject behavior could be compared.

The subjects in our experiments chose strategies maximizing expected gain (MEG), or nearly so (Figure 8.5). Efficiency was defined as the amount of money won relative to the amount of money expected for a subject who used the strategy maximizing expected gain. Subjects' efficiencies were typically above 90% (Trommershäuser *et al.*, 2003a, 2003b, 2005, 2006a; Dean *et al.*, 2007). Subjects chose visuomotor strategies that came close to maximizing gain in a wide variety of simple stimulus configurations, in good agreement with the predictions for the subject maximizing expected gain (Trommershäuser *et al.*, 2003a, 2003b, 2005, 2006a, 2006b; Gepshtein *et al.*, 2007; Stritzke and Trommershäuser, 2007).

The experiments just described focused on spatial movement uncertainty and its consequences for behavior. Time played a role, but only in the time limit imposed on completion of movements to the target area. More recent experiments focus on compensation for temporal uncertainty and, more generally, the allocation of available time. Hudson *et al.* (2008) carried out experiments analogous to Trommershäuser *et al.* (2003a, 2003b), but with subjects rewarded for making touching movements that arrived at a target within a specified time window. If the subject missed the target or arrived outside of the time window, no reward was given. In different experimental conditions the subject could also be penalized for arriving early or late, as summarized in Figure 8.6a. Each of the bars is a time line, and the reward window is colored green. Arriving at times colored in red incurred a penalty, and arriving at times colored neither red nor green incurred no penalty and also no reward. The four reward/penalty conditions in Figure 8.6a were blocked and subjects were informed about the reward/penalty structure; they saw a display similar to the time bars in Figure 8.6a. The challenge for the subject was to compensate for the subject's own temporal movement uncertainty. Figure 8.6b illustrates the one-dimensional computation of expected gain as a function of the temporal aim point selected by the subject. The computation is analogous to that described by Trommershäuser *et al.* (2003a, 2003b), but is now computed in one temporal dimension rather than two spatial dimensions. One difference between the spatial experiments and the temporal experiment of Hudson and colleagues is that temporal movement uncertainty increases with duration of the movement. Hudson *et al.* found that subjects chose temporal aim points in good agreement with those predicted to maximize expected gain in each condition (Figure 8.6c). Moreover, each subject compensated for the increase in timing uncertainty with movements of longer duration.

Dean *et al.* (2007) studied how well subjects traded off speed and accuracy in attempting to hit targets whose value rapidly diminished over time. Once the

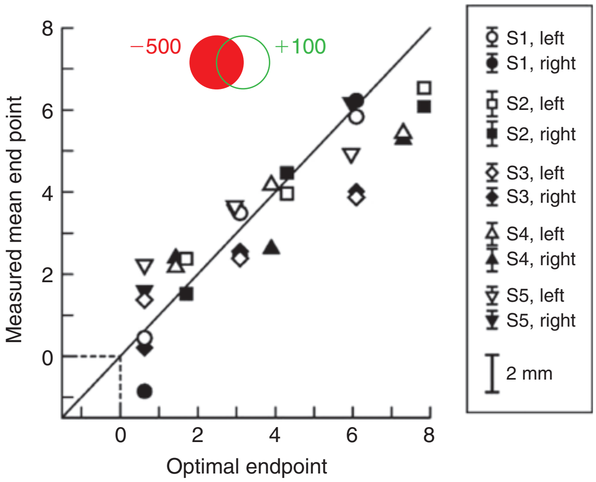


FIGURE 8.5 Movement strategies during rapid movement under risk. Measured mean movement endpoint as a function of optimal endpoint maximizing expected gain (data reprinted from Trommershäuser *et al.*, 2003b) in the task illustrated in Figure 8.3. Five subjects participated. There were six different conditions corresponding to green target circles located to the left (open symbols) or right (filled symbols) of the red penalty circle at each of three different target-penalty distances. The data points fall close to the diagonal identity line, indicating that subjects chose aim points close to those maximizing expected gain based on their individual movement variances.

target appeared, its value decreased from an initial maximum to 0 over 700–1000 ms, depending on condition. The subject received the value associated with the target at the instant that it was hit. If the subject missed the target, of course no reward was given. The challenge for the subject was to determine how to invest a scarce resource (time) by finding the compromise between increasing movement accuracy (by slowing down) and maximizing the value of the target when hit (by speeding up). Dean *et al*. measured subjects' performance in an initial training session, and used these data to estimate each subject's probability of hitting the target as a function of movement duration. This curve is shown in Figure 8.7 (solid curve), together with a dashed line of negative slope that represents the rapidly decreasing value of the target. The product of these two curves is the expected gain as a function of movement duration, and the movement duration associated with maximum expected gain is marked with a circle. The experiment included four conditions that differed in how rapidly the value of the target decreased.

Dean *et al*. found that eight subjects increased or decreased movement duration from condition to condition in accord with predicted MEG duration, but that, overall, subjects were typically 50 ms slow in arriving at the target. This delay was not costly

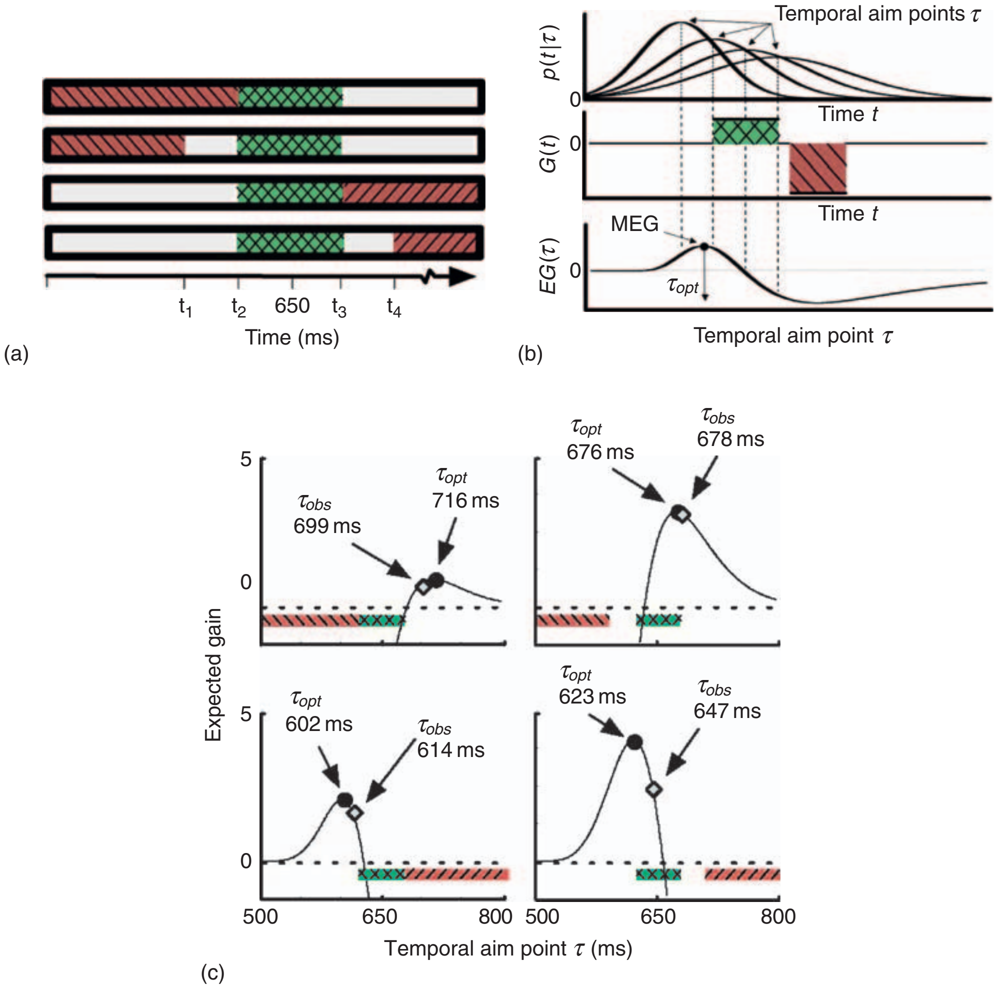


FIGURE 8.6 Movement timing under risk. (a) In this task, subjects reached toward a target and were rewarded if they hit the target arriving within a fixed range of movement times (as indicated by the green portions of these timer bars). In four different conditions, either early or late movements arriving within a second range of movement times were penalized (the red portions of the timer bars). (b) Calculation of expected gain as a function of mean movement duration. Upper panel: distribution of movement times for several mean durations. Note that variance increases with mean duration. Middle panel: expected gain as a function of actual movement duration for one condition. Bottom panel: expected gain varies with mean movement duration; the black circle denotes the mean duration resulting in maximum expected gain. (c) Predictions for one subject; the four panels correspond to the four conditions in (a). Black circles indicate the MEG mean temporal aim point. Diamonds indicate observed mean durations. Across the four conditions and five subjects, mean movement durations were close to those that would result in maximum expected gain (Hudson *et al*., 2008).

(almost all subjects earned 90% or more of their predicted MEG), but it was consistent across subjects and conditions. The authors conjectured that this 50 ms delay might represent a tradeoff between monetary gain and biomechanical cost, as proposed by Trommershäuser *et al*. (2003a, 2003b). In effect, subjects sacrificed some of their potential winnings (about \$0.001/trial) to be able to move slightly less rapidly – a tradeoff of 50 ms per "millidollar."

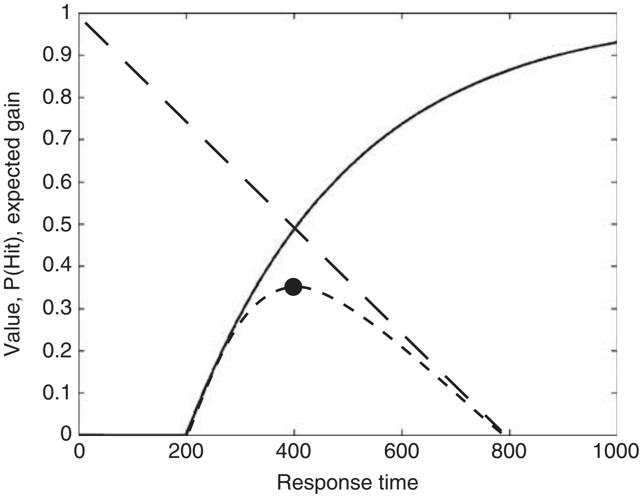


FIGURE 8.7 Calculation of expected gain in the experiment of Dean *et al*. (2007). Subjects reached for a target on a display screen and were rewarded if they hit the target. The value of a hit on the target decreased linearly with movement duration (dashed line). Increased movement duration results in decreased movement endpoint variance, leading to an increased probability of a hit on the target (solid curve). The product of the value and probability of a hit equals expected value (dotted curve). The duration corresponding to maximum expected gain is indicated by the circle.

Battaglia and Schrater (2007) examined human performance in a task where subjects needed time to accurately estimate the location of a target, and also needed time to move to and touch the target to earn rewards. The target was indicated by a random sample of points drawn from a probability density whose centroid was the center of the otherwise invisible target (Figure 8.8a). The points appeared one by one across time, and the longer the subject waited to move the more points would be seen and the more accurate the visual estimate of target location. However, increased viewing time came at the expense of a reduced time allocated to the movement. Moreover, as soon as the subject initiated movement, no further dots marking the target location would appear. There were three experimental conditions with three different probability densities differing in standard deviation ("dot scatter level"). In separate control conditions, they measured the variability of reaches constrained by visual or motor variability alone. Summing these individual variances, they could predict the tradeoff between viewing and movement time that minimized the standard deviation of the endpoint of movement relative to the center of the target (Figure 8.8b). This tradeoff changed with dot scatter level. Their "results suggest that … the brain understands how visual and motor variability depend on time and selects viewing and movement durations to minimize consequent errors" (Battaglia and Schrater, 2007: 6989). These three temporal experiments, taken together, indicate that the

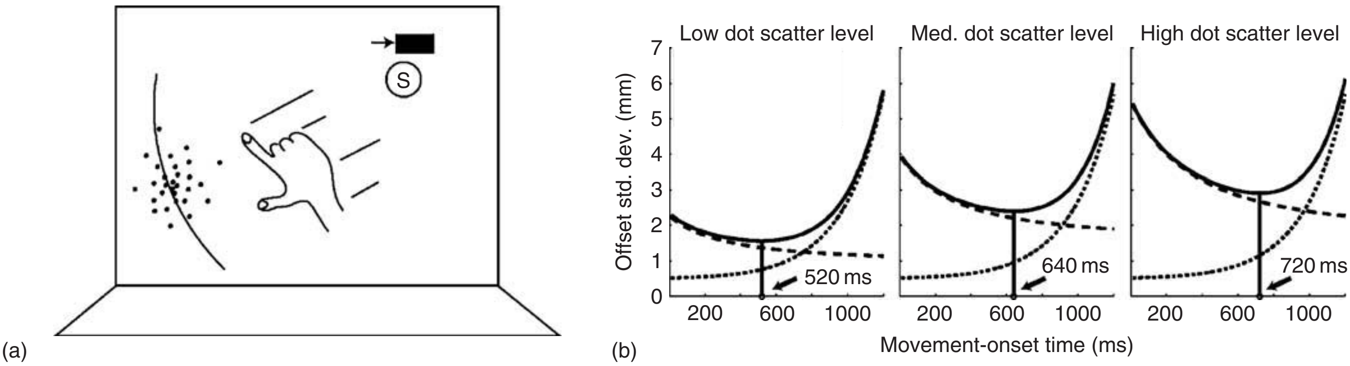


FIGURE 8.8 Task and calculation of the optimal movement strategy in the experiment of Battaglia and Schrater (2007). (a) Targets were indicated by dots drawn from a Gaussian distribution centered on the unseen target. Dots were drawn one by one at a constant rate and no more dots were drawn once the subject began to reach, so that subjects had more information concerning target location the longer they waited to move. "S" indicates the starting point of the reach and the unseen target was located on the circular arc. The black bar indicates the amount of time that had elapsed throughout each trial. (b) Hits on the target were rewarded only if they occurred within 1200 ms of the display of the first dot. The x-axis indicates the time of movement onset. Larger movement-onset times leave less time available for the movement, resulting in greater standard deviation of endpoint location due to motor variability (dotted curve) but smaller standard deviation due to error in visual estimation of target location (dashed curve). The overall standard deviation (relative to the target center) determines the probability of reward. The minimal standard deviation (corresponding to the strategy maximizing expected gain) is indicated. This varied with the standard deviation of the probability from which the dots were sampled ("dot scatter level") as shown in the three panels.

visual system can solve optimization problems involving allocation of a fundamental scarce resource – time.

MOVEMENT PLANNING AND DECISION MAKING

The Importance of Arbitrary Loss Functions

All the models of motor performance discussed above assume that an ideal motor system will adapt to the constraints of a given task by optimizing some measure of cost or expected gain. The three different classes of models examined differ only in the cost to be minimized or the gain to be maximized: biomechanical costs, movement reliability, or economic expected gain. Theories applying optimal control theory to modeling motor behavior (Harris and Wolpert, 1998; Todorov, 2004) as well as applying Bayesian decision theory to modeling adaptive motor behavior (Körding and Wolpert, 2004a), depend crucially on the choice of a fixed and arbitrary loss function. Most implementations of these theories use the mean squared error as the loss function, such that doubling an error quadruples the cost. However, deviations from this assumption have been found for large motor errors, suggesting that the motor system works as a robust estimator by imposing smaller penalties for large errors relative to small than would be expected with a mean-squared-error loss function (Körding and Wolpert, 2004b).

Bayesian inference is currently the most promising candidate for modeling adaptation of the motor system to persistent changes in the movement planner's environment. A movement strategy is a mapping from sensory input[2] v to a movement plan $s(v)$.

The expected gain associated with the choice of strategy $s(v)$ is given by

$$EG(s) = \iiint g(\tau, w)\, p_T(\tau \mid s(v))\, p_V(v \mid w)\, p_W(w)\, dv\, d\tau\, dw, \tag{8.6}$$

where W is the random state of the world with prior distribution $p_W(w)$, V is sensory information about the state of the world with likelihood distribution $p_V(v \mid w)$, and T is the stochastic movement trajectory resulting from the executed movement plan $s(v)$ with distribution $p_T(\tau \mid s(v))$. The term $g(\tau, w)$ specifies the gain resulting from an actual trajectory τ in the actual state of the world w. This gain function can incorporate the gain associated with movement outcome (i.e., whether the actual trajectory τ accomplished the goal specified by world state w), but it can also incorporate the cost of the trajectory itself (i.e., biomechanical costs). Note that gains need not be determined by w and τ; $g(\tau, w)$ need only represent the *expected* gain under those conditions.

Bayesian decision theory makes use of Bayes' rule, which states that one should calculate the *likelihood* of the state of the world, i.e. the probability of the sensory input given the hypothesized state of the world, $p_V(v \mid w)$ The likelihood is integrated with the *prior* $p_W(w)$, which reflects the subject's belief about the particular state of the world before the sensory input is received. The prior may reflect how probable it is that objects in the world have a particular size or are oriented in a particular way. By multiplying the prior and the likelihood and normalizing (scaling so that the probabilities over all possible states sum to 1), we can estimate the probability of the state given the sensory input $p(w \mid v)$, termed the *posterior* of the state. This posterior could then become the new prior belief, and could be further updated based on later sensory input (see, for example, review by Bays and Wolpert, 2006).

Motor adaptation in agreement with Bayesian inference has recently been demonstrated in a study in which subjects made online corrections to a reaching movement based on momentary visual feedback of hand position (Körding and Wolpert, 2004a). Visual feedback, presented midway through the movement, was displaced laterally by a distance that varied randomly from trial to trial. The range of displacements experienced over the course of many trials formed a prior probability distribution. According to the Bayesian model, the prior distribution was combined with the feedback on a given trial to provide the best estimate of the position error. For movements in which very precise visual feedback was given, the prior distribution of displacements had little influence on the estimate of hand position. However, when the visual feedback was artificially blurred, the state estimate became increasingly biased towards the mean of the prior distribution, as predicted by a Bayesian model. In a later study, Tassinari *et al*. (2006) asked subjects to point at targets indicated by unreliable visual information, where targets were drawn from a prior distribution. Subjects had previously been trained on the prior distribution. Subjects displayed a similar shift from the position indicated by the visual information toward the mean of the prior distribution, although the variation of shift amounts across experimental conditions was smaller than of the predicted shifts that would maximize expected gain.

A similar result was found in a task in which subjects' movements were disturbed by force pulses of

---

[2] We follow the convention that random variables are in upper case, e.g. X, while the corresponding specific values that those variables can take on are in lower case, e.g. $p(x)$.

varying amplitude (Körding *et al.*, 2004). The prior probability distribution of the strength of the force perturbation could be estimated by the subject as the distribution of forces experienced over the course of the experiment. Subjects experienced two force pulses during a single movement, and were told that they were of equal strength. Thus, to resist the second force pulse, subjects had to predict its strength by combining their estimate of the first pulse with their current estimate of the mean of the prior distribution. Subjects adapted their behavior in a manner consistent with the predictions of Bayesian integration of a noisy estimate of the force of the first pulse with the prior. The contribution of the prior depended appropriately on the prior information.

Learning vs Computing

Surprisingly, subjects do not show a trend of gradually approaching maximum expected gain during these movement-planning tasks under risk. However, these effects may be explained by the specific design of the studies (e.g., Trommershäuser *et al.*, 2003a, 2003b, 2005; Dean *et al.*, 2007). Before the "decision-making" phase of the experiment, subjects practiced the speeded motor task extensively by simply touching green targets. During this initial training period, the experimenter monitored their motor performance until subjects' reaction times stabilized to the time constraints of the tasks and the experimenter could measure each subject's residual motor variability. Following training, subjects learned about the gains and losses assigned to each region and were asked to try to earn as much money as they could by hitting the green circle and trying to avoid hitting the penalty region. Subjects were not explicitly instructed to work out a motor strategy that took into account the spatial locations of reward and penalty regions and the magnitude of penalty and reward, but their highly efficient performance indicates that they did so from the first trial in which rewards and penalties were specified. To summarize, in the design of Trommershäuser *et al.* (2003a, 2003b) and later work by Dean *et al.* (2007), subjects were first trained to be "motor experts" in speeded pointing towards single targets on the screen. Only then were they confronted with a task involving tradeoffs between possible rewards and penalties. As Trommershäuser *et al.* (2003a, 2003b, 2005) reported, there were no obvious trends in subjects' aim points that would suggest that subjects were modifying their

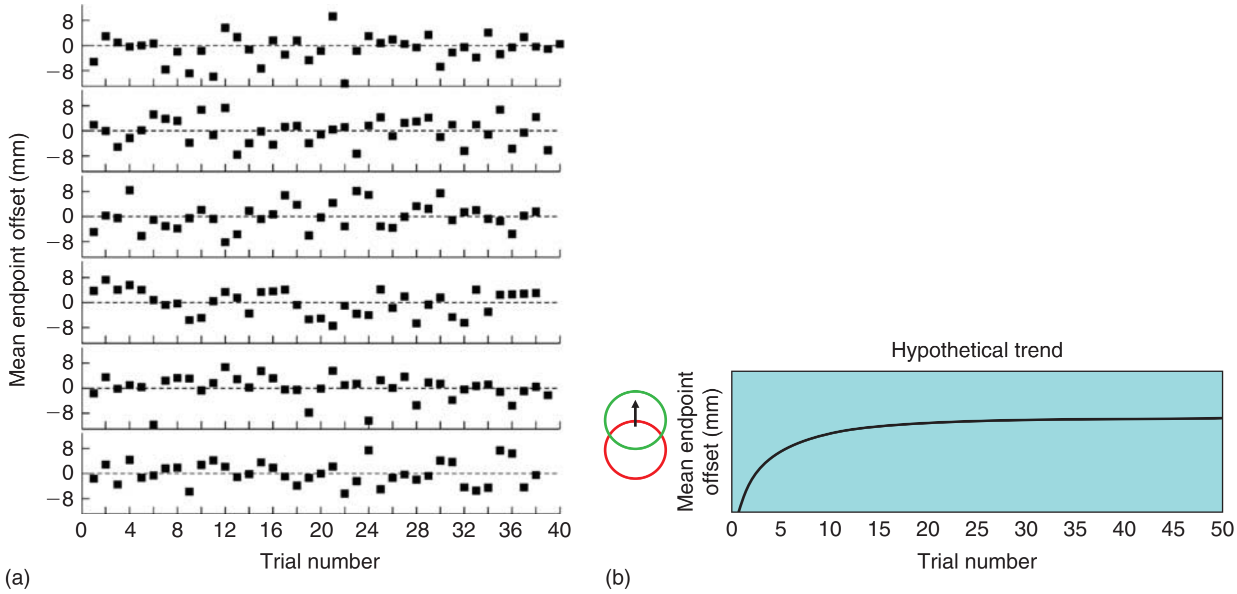


FIGURE 8.9 Consistency of pointing strategy across trials. (a) Trial-by-trial deviation of movement endpoint from mean movement endpoint as a function of trial number after introduction of rewards and penalties; the six different lines correspond to the six different spatial conditions of target and penalty offset (data replotted from Figure 7, Trommershäuser *et al.*, 2003a). (b) Trend of a hypothetical learning model in which a subject changes motor strategy gradually in response to rewards and penalties incurred. The subject initially aims at the center of the green circle. Before the subject's first trial in the decision-making phase of the experiment, the subject is instructed that red circles carry penalties and green circles carry rewards. Subjects may approach the optimal aim point maximizing expected gain by slowly shifting the aim point away from the center of the green circle until the winnings match the maximum expected gain. However, the data shown in (a) do not exhibit this trend and do not support such a learning model.

decision-making strategy in response to their experience with the decision-making task (Figure 8.9a).

To see how unusual this finding is, consider applying a hypothetical learning model that changes motor strategy only gradually in response to rewards and penalties incurred (see, for example, Sutton and Barto, 1998; Dayan and Balleine, 2002; Daw and Doya, 2006). In the training part of our experiments, subjects learned to aim toward the center of the green circle. After the training, and before the subject's first trial in the decision-making phase of the experiment, the subject was instructed that red circles carry penalties and green circles carry rewards. What should the subject do on the first trial of the decision-making phase of the experiment? In the absence of any reward or penalty, a learning model based on reward and penalty would predict that the subject should aim at the center of the green circle, just as in the training trials. The subject would then gradually alter his/her motor strategy in response to the rewards and penalties incurred until the final motor strategy approximated the strategy maximizing expected gain (Figure 8.9b).

However, examination of the initial trials of the decision phase of the experiment (Figure 8.9a) suggests that subjects immediately changed their movement strategy from that used in training to that required to optimally trade off the probabilities of hitting the reward and penalty regions. This apparent lack of learning is of great interest in that it suggests that, while subjects certainly learned to carry out the motor task in the training phases of these experiments, and learned their own motor uncertainty, they seemed not to need further experience with the decision-making task to perform as well as they did.

Motor and Perceptual Decisions

Comparing subjects' behavior with the predictions of an ideal performance model allows us to explore the limits of information processing during goal-directed behavior. In some of our previous work, we have varied basic parameters of the model asking how human behavior deviates from optimal behavior maximizing expected gain once the integration of sensory, motor, and reward information becomes too difficult or too time costly. Subjects do not always come close to maximizing expected gain in movement-planning tasks under risk. We find that subjects are able to optimally plan their movements, as long as full information about the stimulus configuration and the assigned rewards is provided prior to movement onset (Trommershäuser *et al.*, 2006b). Subjects fail to select strategies that maximize expected gain in motor tasks similar to that of Trommershäuser *et al.* (2003a, 2003b) when there is a reward region and more than one penalty region (Wu *et al.*, 2006), when target and penalty circles are reduced in contrast and blurred (Ma-Wyatt *et al.*, 2006), and when rewards and penalties are awarded according to a stochastic payoff rule (Maloney *et al.*, 2007). Moreover, in unspeeded *visual* tasks analogous to those of Trommershäuser *et al.* (2003a, 2003b), subjects fail to compensate for trial-to-trial variation in uncertainty (Landy *et al.*, 2007). Thus, while there is a collection of speeded motor tasks under risk where performance is remarkably efficient, we cannot simply generalize these results to a broader claim that performance in any perceptual or motor task under risk would be "near-optimal."

Movement Under Risk, Decision Making Under Risk

In planning movement in our movement tasks under risk, our subjects are effectively choosing from among a set of many possible lotteries. To see why, consider a trial in which hits on the target and penalty yield gains of +1 and −5 cents, respectively (Figure 8.4a). In executing the movement, the subject chooses a strategy S which we've identified with the mean endpoint $(\bar{x}, \bar{y})$. The choice of strategy fixes the probability $P(R_i \mid S)$ of hitting the target region, the penalty region, the region where target and penalty overlap, and the background, and hence of being awarded the gains G_i associated with each region. In the decision-making literature, this combination of event probabilities $P(R_i \mid S)$ and associated gains G_i is called a "lottery" $L(S)$,

$$\begin{aligned} L(S) = (&P(R_1, S), G_1; P(R_2, S), G_2; \\ &P(R_3, S), G_3; P(R_4, S), G_4). \end{aligned} \tag{8.7}$$

An alternative movement strategy S' corresponds to a second lottery

$$\begin{aligned} L(S') = (&P(R_1, S'), G_1; P(R_2, S'), G_2; \\ &P(R_3, S'), G_3; P(R_4, S'), G_4). \end{aligned} \tag{8.8}$$

As illustrated in Figure 8.4b–c, every mean endpoint results in a lottery with a corresponding expected gain, i.e., the expected number of points a subject will earn, on average, having "aimed" at $(\bar{x}, \bar{y})$. However, there are many other possible mean endpoints and corresponding lotteries, each with its associated expected gain. By choosing among all these possible strategies, the subjects in our experiments effectively select among the many possible lotteries.

The results of our experiments indicate that subjects choose strategies maximizing expected gain, or nearly so. Efficiency was defined as the amount of money won relative to the amount expected using an optimal strategy maximizing expected gain. Subjects' efficiencies were typically above 90% (Trommershäuser *et al.*, 2003a, 2003b).

In contrast to the highly efficient visuo-motor strategies observed during visuo-motor tasks under risk, human decision makers in decision making under risk typically fail to maximize expected gain. Expected utility theory (Bernoulli, 1738/1954; von Neumann and Morgenstern, 1944) is based on the assumption that subjects assign numerical utilities to outcomes and maximize expected utility. An evident advantage of the utility hypothesis is that a wide range of consequences (e.g., biomechanical costs and money) can be measured in identical units and it becomes meaningful to seek optimal tradeoffs among them. The model we presented assumes that this is the case.

When outcomes are specified as monetary rewards (as in our experiments), utility can be a non-linear function of monetary rewards and can also depend on subjects' current "wealth," and consequently a subject maximizing expected utility would fail to maximize expected gain. Bernoulli (1738/1954) originally suggested that this utility function is a concave function of value which increases quickly and then flattens out for larger values, and that the shape of the utility function could explain observed risk aversion in human decision making. Non-linear utility functions applied to gains and losses have been employed in describing human performance in a variety of economic decision tasks (Kahneman *et al.*, 1982; Bell *et al.*, 1988; Kahneman and Tversky, 2000). There is nothing inherently suboptimal or erroneous in maximizing expected utility, however it is defined, since by definition utility is what the organism seeks to maximize.

Moreover, several factors may have contributed to subjects' tendency to maximize expected gain found in our studies (Trommershäuser *et al.*, 2003a, 2003b, 2005; Dean *et al.*, 2007). In the motor task the subject makes a long series of choices and, over the course of the experiment, the accumulated winnings increase. On the other hand, subjects in economic decision making experiments typically make a single, "one-shot" choice between a discrete set of lotteries. Indeed, when economic decision makers are faced with a series of decisions, they tend to move closer to maximum expected gain (Redelmeier and Tversky, 1992; Wakker *et al.*, 1997; "the house money effect," Thaler and Johnson, 1990). Studies of risky choice find that subjects are closer to maximizing expected gain for small stakes (Camerer, 1992; Holt and Laury, 2002), and when subjects receive feedback over the course of the experiment or have prior experience with the task (Hertwig *et al.*, 2004). All of these factors would tend to "linearize" subjects' utility functions and to move them toward maximizing expected gain.

However, human performance in decision-making tasks is markedly suboptimal by other relevant criteria. Deviations from expected utility theory include a tendency to change one's decision based on whether the lottery was described in terms of losses or gains, due to an exaggerated aversion to losses (Kahneman and Tversky, 1979) and a tendency to exaggerate small probabilities (Allais, 1953; Attneave, 1953; Lichtenstein *et al.*, 1978; Tversky and Kahneman, 1992). This exaggeration of the frequency of low-frequency events is observed in many, but not all, decision-making studies (Sedlmeier *et al.*, 1998). These distortions of probability would, if present in movement planning, be particularly problematic. The strategies maximizing expected gain in many of the motor tasks above involve small probabilities of large losses (Figure 8.4) and exaggerated-aversion of losses and overweighting of small probabilities would likely impair performance.

The contrast between success in "movement-planning under risk" and decision making under risk is heightened by the realization that, in cognitive decision making under risk, subjects are told the exact probabilities of outcomes and thus have perfect knowledge of how their choice of strategy changes the probability of attaining each outcome. The knowledge of probabilities in equivalent motor tasks is never communicated explicitly, but is acquired across a few hundred training trials, and thus can equal but never exceed the knowledge available under cognitive decision making under risk. The results of our experiments imply that subjects are able to learn their own motor uncertainties very well (Trommershäuser *et al.*, 2005; Gepshtein *et al.*, 2007; see also Baddeley *et al.*, 2003). These results suggest that humans are able to estimate the uncertainties associated with sensory and motor noise and make use of this knowledge to improve their performance.

In summary, the results of our work indicate that movement planning shares the same formal structure as perceptual decision making and economic decision making. Subjects in movement tasks are generally found to be very good at choosing motor strategies that come close to maximizing expected gain. In contrast, subjects in economic decision making and perceptual estimation tasks typically fail to maximize expected gain. Moreover, the sources of uncertainty in motor tasks are endogenous: they reflect the organism's own uncertainty in planning and executing movement while, in contrast, uncertainty in economic tasks is typically imposed by the experimenter. Thus probabilistic information from cognition, perception, and movement has different origins, and it would be of interest for future work to compare the neural circuits underlying the representation of probability in movement and economic decision making.

NEURAL CORRELATES OF MOTOR AND COGNITIVE DECISIONS

We finally summarize recent experimental work directed at understanding the neural coding of motor and cognitive decisions. Most of the current evidence results from electrophysiological recordings in monkeys measuring single-cell activity during binary-decision tasks in response to manipulations of reward and reward uncertainty (see, for example, Sugrue *et al.* (2005) for a review).

Following Herrnstein's (1961) pioneering behavioral work, electrophysiological studies typically employ a paradigm in which a monkey chooses between two alternative responses that may differ with respect to the sensory information available on each trial, the prior odds, and the outcome assigned to each response alternative. These experiments yield insight into how sensory information is integrated with reward information accumulated across previous trials. Reward is typically manipulated by assigning variable amounts of juice to different color-coded response alternatives (Platt and Glimcher, 1999; Sugrue *et al.*, 2004; see also Chapters 6, 30, and 32). When rewards were assigned stochastically, the monkey's choices appeared to be based on an estimate of the probability of reward gained by sampling over the last few trials (Sugrue *et al.*, 2004). These results indicate that the brain quickly reaches a decision based on the reward history of the last few trials.

Single cell activity in response to stochastic variations of reward has been found in ventral midbrain areas (Fiorillo *et al.*, 2003). These dopamine neurons' phasic activity correlates with the so-called prediction error, i.e. with the difference between actual and expected reward (Schultz *et al.*, 1997; Morris *et al.*, 2004; see also Part 4 of this volume, pp. 321–416), and modulating that activity affected choice (Pessiglione *et al.*, 2006). However, these same neurons also produced a tonic response that was highest in conditions of highest risk, i.e. in trials in which the probability of receiving the reward was 0.5. The behavioral relevance of this midbrain dopaminergic single cell activity recorded in response to changes in reward probability remains controversial (Bayer and Glimcher, 2005; Morris *et al.*, 2004; Niv *et al.*, 2006).

Using fMRI techniques in humans, a variety of subcortical and cortical areas have been implicated in the coding of decision variables such as expected gain, probability of reward, risk, and ambiguity. Most studies employ a visual representation of a gambling task and, after a delay of several seconds, subjects are instructed to choose between pairs of options by key-press. Brain activity is monitored during the delay period and correlated with various decision variables (see, for example, Glimcher and Rustichini, 2004; O'Doherty, 2004; Rorie and Newsome, 2005; Trepel *et al.*, 2005; Daw and Doya, 2006; Montague *et al.*, 2006). Neural activity may be correlated with gain or loss of a potential reward, the probability of reward, their product (expected gain) or risk (the variance of gain). A number of studies suggest that reward value is encoded in the striatum and portions of prefrontal cortex (PFC) and orbitofrontal cortex (OFC) (O'Doherty, 2004; Knutson *et al.*, 2005; Tanaka *et al.*, 2005; Daw *et al.*, 2006; Tom *et al.*, 2007).

In humans, reward-prediction error signals are generally found to be localized to the striatum, although they are also seen in OFC and amygdala (O'Doherty, 2004; Daw *et al.*, 2006; Pessiglione *et al.*, 2006; Yacubian *et al.*, 2006). It has been difficult to disentangle probability of reward from expected gain, and most studies find responses correlated with expected gain in the striatum, OFC and medial PFC (Delgado *et al.*, 2005; Hsu *et al.*, 2005; Knutson *et al.*, 2005; Daw *et al.*, 2006; Preuschoff *et al.*, 2006). Several of these studies see an increase in activity with increasing expected value whether or not the outcome is an expected loss or gain. On the other hand, Yacubian *et al.* (2006) suggest that while expected gains are encoded in the striatum, expected losses result in responses in the amygdala, perhaps also associated with negative emotion. This is supported by the finding that decisions consistent with framing effects are correlated with increased response in the amygdala. PFC response is higher in subjects that have less of a framing effect (De Martino *et al.*, 2006), suggesting a requirement for cognitive control to suppress this cognitive bias.

An important distinction in these gambling tasks is between risk (in which the probabilities of the outcomes are known precisely) and ambiguity (when they are not). Responses correlated with risk have been found in anterior insula, OFC and striatum (Preuschoff *et al.*, 2006), in the ventral striatum and anterior insula (Kuhnen and Knutson, 2005), as well as in dorsolateral PFC and posterior parietal cortex (Huettel *et al.*, 2005).

Responses correlated with the ambiguity of a decision have been found in the posterior part of the inferior frontal sulcus (Huettel *et al.*, 2006), OFC, amygdala, and dorsomedial PFC, along with a negative correlation with responses in the striatum (Hsu *et al*., 2005). Subjects with a preference for ambiguity over risk show stronger responses in lateral PFC, while subjects with a preference for risk over ambiguity show stronger responses in posterior parietal cortex (Huettel *et al*., 2006). A reward received immediately is generally valued more than one that will be delayed – a phenomenon known as temporal discounting. If a reward will be received immediately, a variety of brain areas respond, including striatum and OFC; however, the inclusion of a potentially delayed reward recruits other areas, including portions of the PFC (McClure *et al*., 2004; Tanaka *et al*., 2005; Glimcher *et al*., 2007), suggesting the need for cognitive control for choices involving delayed gratification.

Little is known about the neural coding of errors in pure motor tasks. Comparing errors in movement completion (induced by target displacements) with kinematic errors (induced by novel visual feedback) and with dynamic errors (induced by the application of force fields) showed increased cerebellar activity both for kinematic and for dynamic errors (Diedrichsen *et al*., 2005). Target errors, but not execution errors, activated the posterior superior parietal lobule and the striatum. In contrast, execution errors produced strong adaptive responses that specifically activated anterior aspects of the parietal cortex and the dorsal premotor cortex. Overall, structures involved in the correction of errors attributable to mis-estimation of dynamics were generally a subset of the neural areas involved in correction of movement errors attributable to mis-estimation of kinematics.

CONCLUSION

We have presented results from a variety of different approaches directed at understanding the processes underlying decision making in motor tasks. The results presented here indicate that movement planning shares the same formal structure as economic decision making. Subjects in movement tasks are generally found to be very good at choosing motor strategies that come close to maximizing expected gain. In contrast, subjects in economic decision making typically fail to maximize expected gain. Moreover, the sources of uncertainty in motor tasks are endogenous; they reflect the organism's own uncertainty in planning movement while, in contrast, uncertainty in economic tasks is typically imposed by the experimenter. Thus, probabilistic information from cognition, perception, and movement has different origins.

In economic decision tasks, feedback regarding outcomes typically reduces biases and misperceptions in the representation of probability estimates, moving behavior closer to strategies maximizing expected gain. We emphasize that in movement-planning under risk subjects' performance is initially close to optimal performance, maximizing expected gain, and does not appear to change with feedback. Movement planning is well described by simple models that maximize expected gain while there is no single model of economic decision making that captures all of the complexity of human behavior. Careful study of the neural circuitry underlying decision making in the form of movement could lead to a better understanding of how the brain gathers information to make decisions and transforms them into movement.

Acknowledgments

We thank Paul Glimcher for his comments on an earlier draft of this manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (Emmy-Noether-Programme; grant TR 528/1–3) and the National Institutes of Health (grant EY08266).

References

Alexander, R.M. (1997). A minimum energy cost hypothesis for human arm trajectories. *Biol. Cybern.* 76, 97–105.

Allais, M. (1953). Le comportment de l'homme rationnel devant la risque: critique des postulats et axiomes de l'école Américaine. *Econometrica* 21, 503–546.

Attneave, F. (1953). Psychological probability as a function of experienced frequency. *J. Exp. Psychol.* 46, 81–86.

Baddeley, R.J., Ingram, H.A., and Miall, R.C. (2003). System identification applied to a visuomotor task: near-optimal human performance in a noisy changing task. *J. Neurosci.* 7, 3066–3075.

Battaglia, P.W. and Schrater, P.R. (2007). Humans trade off viewing time and movement duration to improve visuomotor accuracy in a fast reaching task. *J. Neurosci.* 27, 6984–6994.

Bayer, H.M. and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141.

Bays, P.M. and Wolpert, D.M. (2006). Computational principles of sensorimotor control that minimize uncertainty and variability. *J. Physiol.* 578, 387–396.

Bell, D.E., Raiffa, H., and Tversky, A. (eds) (1988). *Decision Making: Descriptive, Normative and Prescriptive Interactions*. Cambridge: Cambridge University Press.

Bernoulli, D. (1738/1954). Exposition of a new theory on the measurement of risk [*Comentarii Academiae Scientiarum Imperialis Petropolitanae*]. Translation published in *Econometrica*, 22, 23–36.

Burdet, E., Osu, R., Franklin, D.W. *et al*. (2001). The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature* 414, 446–449.

Camerer, C.F. (1992). The rationality of prices and volume in experimental market. *Organ. Behav. Hum. Dec. Proc.* 51, 237–272.

Cuijpers, R.H., Smeets, J.B.J., and Brenner, E. (2004). On the relation between object shape and grasping kinematics. *J. Neurophysiol.* 91, 2598–2606.

Daw, N.D. and Doya, K. (2006). The computational neurobiology of learning and reward. *Curr. Opin. Neurobiol.* 16, 199–204.

Daw, N.D., O'Doherty, J.P., Dayan, P. *et al*. (2006). Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.

Dayan, P. and Balleine, B.W. (2002). Reward, motivation and reinforcement learning. *Neuron* 36, 285–298.

Dean, M., Wu, S.-W., and Maloney, L.T. (2007). Trading off speed and accuracy in rapid, goal-directed movements. *J. Vision,* 7, 1–12.

Delgado, M.R., Miller, M.M., Inati, S., and Phelps, E.A. (2005). An fMRI study of reward-related probability learning. *NeuroImage* 24, 862–873.

De Martino, B., Kumaran, D., Seymour, B., and Dolan, R.J. (2006). Frames, biases, and rational decision-making in the human brain. *Science* 313, 684–687.

Diedrichsen, J., Hashambhoy, Y., Rane, T., and Shadmehr, R. (2005). Neural correlates of reach errors. *J. Neurosci.* 25, 9919–9931.

Dornay, M., Uno, Y., Kawato, M., and Suzuki, R. (1996). Minimum muscle-tension change trajectories predicted by using a 17-muscle model of the monkey's arm. *J. Mot. Behav.* 2, 83–100.

Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902.

Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.* 47, 381–391.

Flash, T. and Hogan, N. (1985). The coordination of arm movements: an experimentally confirmed mathematical model. *J. Neurosci.* 5, 1688–1703.

Franklin, D.W., Liaw, G., Milner, T.E. *et al*. (2007). Endpoint stiffness of the arm is directionally tuned to instability in the environment. *J. Neurosci.* 27, 7705–7716.

Gepshtein, S., Seydell, A., and Trommershäuser, J. (2007). Optimality of human movement under natural variations of visual-motor uncertainty. *J. Vision* 7, 1–18.

Glimcher, P.W. and Rustichini, A. (2004). Neuroeconomics: the consilience of brain and decision. *Science* 306, 447–452.

Glimcher, P.W., Kable, J., and Louie, K. (2007). Neuroeconomic studies of impulsivity: now or just as soon as possible? *Am. Econ. Rev.* 97, 142–147.

Hamilton, A.F.C. and Wolpert, D.M. (2002). Controlling the statistics of action: obstacle avoidance. *J. Neurophysiol.* 87, 2434–2440.

Harris, C.M. and Wolpert, D.M. (1998). Signal-dependent noise determines motor planning. *Nature* 394, 780–784.

Herrnstein, R.J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* 4, 267–272.

Hertwig, R., Barron, G., Weber, E.U., and Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychol. Sci.* 15, 534–539.

Holt, C. and Laury, S. (2002). Risk aversion and incentive effects. *Am. Econ. Rev.* 92, 1644–1655.

Hsu, M., Bhatt, M., Adolphs, R. *et al*. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310, 1680–1683.

Hudson, T.E., Maloney, L.T., and Landy, M.S (2008). Optimal movement timing with temporally asymmetric penalties and rewards. PLoS Computational Biology. In press.

Huettel, S.A., Song, A.W., and McCarthy, G. (2005). Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *J. Neurosci.* 25, 3304–3311.

Huettel, S.A., Stowe, C.J., Gordon, E.M. *et al*. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron* 49, 765–775.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–292.

Kahneman, D. and Tversky, A. (eds) (2000). *Choices, Values and Frames*. New York, NY: Cambridge University Press.

Kahneman, D., Slovic, P., and Tversky, A. (eds) (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Kaminsky, T. and Gentile, A.M. (1986). Joint control strategies and hand trajectories in multijoint pointing movements. *J. Mot. Behav.* 18, 261–278.

Knutson, B., Taylor, J., Kaufman, M. *et al*. (2005). Distributed neural representation of expected value. *J. Neurosci.* 25, 4806–4812.

Körding, K.P. and Wolpert, D.M. (2004a). Bayesian integration in sensorimotor learning. *Nature* 427, 244–247.

Körding, K.P. and Wolpert, D.M. (2004b). The loss function of sensorimotor learning. *Proc. Natl Acad. Sci. USA* 101, 9839–9842.

Körding, K.P., Ku, S.P., and Wolpert, D.M. (2004). Bayesian integration in force estimation. *J. Neurophysiol.* 92, 3161–3165.

Kuhnen, C.M. and Knutson, B. (2005). The neural basis of financial risk taking. *Neuron* 47, 763–770.

Landy, M.S., Goutcher, R., Trommershäuser, J., and Mamassian, P. (2007). Visual estimation under risk. *J. Vision* 7(4), 1–15.

Lichtenstein, S., Slovic, P., Fischhoff, B. *et al*. (1978). Judged frequency of lethal events. *J. Exp. Psychol. Hum. Learn.* 4, 551–578.

Maloney, L.T., Trommershäuser, J., and Landy, M.S. (2007). Questions without words: a comparison between decision making under risk and movement planning under risk. In: W. Gray (ed.), *Integrated Models of Cognitive Systems*. New York, NY: Oxford University Press, pp. 297–315.

Ma-Wyatt, A., Stritzke, M., and Trommershäuser, J. (2006). eye–hand coordination for rapid pointing feed back can be used to alter. *J. Vis*, 6, 920a.

McClure, S.M., Laibson, D.I., Lowenstein, G., and Cohen, J.D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507.

Meyer, D.E., Abrams, R.A., Kornblum, S. *et al*. (1988). Optimality in human motor performance: ideal control of rapid aimed movements. *Psychol. Rev.* 95, 340–370.

Montague, P.R., King-Casas, B., and Cohen, J.D. (2006). Imaging valuation models in human choice. *Annu. Rev. Neurosci.* 29, 417–448.

Morris, G., Arkadir, D., Nevet, A. *et al*. (2004). Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* 43, 133–143.

Niv, Y., Daw, N.D., and Dayan, P. (2006). Choice value. *Nature Neurosci.* 9, 987–988.

O'Doherty, J.P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* 14, 769–776.

Pessiglione, M., Seymour, B., Flandin, G. *et al*. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442, 1042–1045.

Platt, M.L. and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.

Preuschoff, K., Bossaerts, P., and Quartz, S.R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51, 381–390.

Redelmeier, D.A. and Tversky, A. (1992). On the framing of multiple prospects. *Psychol. Sci.* 3, 191–193.

Rorie, A.E. and Newsome, W.T. (2005). A general mechanism for decision-making in the human brain? *Trends Cogn. Sci.* 9, 41–43.

Sabes, P.N. and Jordan, M.I. (1997). Obstacle avoidance and a perturbation sensitivity model for motor planning. *J. Neurosci.* 17, 7119–7128.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Sedlmeier, P., Hertwig, R., and Gigerenzer, G. (1998). Are judgments of the positional frequencies of letters systematically biased due to availability? *J. Exp. Psychol. Learn. Mem. Cogn.* 24, 754–770.

Soechting, J.F. and Lacquaniti, F. (1981). Invariant characteristics of a pointing movement in man. *J. Neurosci.* 1, 710–720.

Soechting, J.F., Buneo, C.A., Herrmann, U., and Flanders, M. (1995). Moving effortlessly in three dimensions: does Donders' Law apply to arm movement? *J. Neurosci.* 15, 6271–6280.

Stritzke, M. and Trommershäuser, J. (2007). Rapid visual localization during manual pointing under risk. *Vision Res.* 47, 2000–2009.

Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.

Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2005). Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* 6, 363–375.

Sutton, R.S. and Barto, A.G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Tanaka, S.C., Doya, K., Okada, G. *et al*. (2005). Prediction of immediate and future rewards differentially recruits cortico–basal ganglia loops. *Nat. Neurosci.* 7, 887–893.

Tassinari, H., Hudson, T.E., and Landy, M.S. (2006). Combining priors and noisy visual cues in a rapid pointing task. *J. Neurosci.* 26, 10154–10163.

Thaler, R. and Johnson, E.J. (1990). Gambling with the house money and trying to break even: the effects of prior outcomes on risky choice. *Management Sci.* 36, 643–660.

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nat. Neurosci.* 9, 907–915.

Tom, S.M., Fox, C.R., Trepel, C., and Poldrack, R.A. (2007). The neural basis of loss aversion in decision-making under risk. *Science* 315, 515–518.

Trepel, C., Fox, C.R., and Poldrack, R.A. (2005). Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk. *Brain Res. Cogn. Brain Res.* 23, 34–50.

Trommershäuser, J., Maloney, L.T., and Landy, M.S. (2003a). Statistical decision theory and tradeoffs in the control of motor response. Spat. Vis. 16, 255–275.

Trommershäuser, J., Maloney, L.T., and Landy, M.S. (2003b). Statistical decision theory and the selection of rapid, goal-directed movements. *J. Opt. Soc. Am. A* 20, 1419–1433.

Trommershäuser, J., Gepshtein, S.G., Maloney, L.T. *et al*. (2005). Optimal compensation for changes in task-relevant movement variability. *J. Neurosci.* 25, 7169–7178.

Trommershäuser, J., Landy, M.S., and Maloney, L.T. (2006a). Humans rapidly estimate expected gain in movement planning. *Psychol. Sci.* 17, 981–988.

Trommershäuser, J., Mattis, J., Maloney, L.T., and Landy, M.S. (2006b). Limits to human movement planning with delayed and unpredictable onset of needed information. *Exp. Brain Res.* 175, 276–284.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. *J. Risk Uncertain.* 5, 297–323.

Uno, Y., Kawato, M., and Suzuki, R. (1989). Formation and control of optimal trajectory in human multijoint arm movement: minimum torque-change model. *Biol. Cybern.* 61, 89–101.

von Neumann, J. and Morgenstern, O. (1944). *The Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Wakker, P.P., Thaler, R.H., and Tversky, A. (1997). Probabilistic insurance. *J. Risk Uncertain.* 15, 7–28.

Wu, S.-W., Trommershäuser, J., Maloney, L.T., and Landy, M.S. (2006). Limits to human movement planning in tasks with asymmetric value landscapes. *J. Vision* 5, 53–63.

Yacubian, J., Gläscher, J., Schroeder, K. *et al*. (2006). Dissociable systems for expected gain- and loss-related value predictions and errors of prediction in the human brain. *J. Neurosci.* 26, 9530–9537.

PART II

BEHAVIORAL ECONOMICS AND THE BRAIN

CHAPTER

9

The Psychology and Neurobiology of Judgment and Decision Making: What's in it for Economists?

B. Douglas Bernheim

OUTLINE

INTRODUCTION

The last few years have witnessed impressive progress toward understanding the neurobiology of decision making. Many participants in this growing field, as well as interested observers, hope that neuroeconomics will eventually make foundational contributions to the various traditional fields from which it emerged, including economics, psychology, and artificial intelligence. My purpose here is to evaluate its potential contributions to economics.

Some would argue that any aspect of economic decision making is definitionally an aspect of economics. According to that view, neuroeconomics necessarily contributes to economics by expanding the set of empirical questions that economists can address. I will avoid such semantic quibbles. My interest here is in assessing whether, in time, neuroeconomics is likely to shed useful light on traditional economic questions. While the scope of traditional economics is difficult to define with precision, I am content with an operational definition, based on the collection of questions and issues currently discussed in standard economic textbooks and leading professional journals.

The potential importance of neuroeconomics for economic inquiry has already been the subject of much debate. For example, an optimistic assessment appeared in a paper titled "Neuroeconomics: Why Economics Needs Brains," by Colin Camerer, George Loewenstein, and Drazen Prelec (2004)[1]. Subsequently, Faruk Gul and Wolfgang Pesendorfer (2008) penned a broad critique of neuroeconomics, titled "The Case for Mindless Economics," which expressed deeply rooted skepticism. My assessment lies between those extremes. I caution against dismissing the entire field merely because current technology is limited, or because some of the early claims concerning its potential contributions to standard economics were excessive and/or poorly articulated. However, because I share many of the conceptual concerns raised by Gul and Pesendorfer, I also see a pressing need for a sober and systematic articulation of the field's relevance. Such an articulation would ideally identify standard economic questions of broad interest (e.g., how taxes affect saving), and outline conceivable research agendas based on actual or potential technologies that could lead to specific, useful insights of direct relevance to those questions. Vague assertions that a deeper understanding of decision-making processes will lead to better models of choice will not suffice to convince the skeptics.

In Bernheim (2008), I have attempted to identify and articulate the specific ways in which neuroeconomics might contribute to mainstream economics, as well as the limitations of those potential contributions. This chapter briefly summarizes both my reservations and my reasons for guarded optimism. Due to space constraints, it touches only lightly on many important issues; readers are referred to the longer version for a more comprehensive and detailed discussion. Perhaps most significantly, I focus here exclusively on positive economics, as does nearly all existing research on neuroeconomics[2]. For the reasons discussed in Bernheim (2008), the possible applications of neuroeconomics to normative economic analysis are intriguing and largely unexplored[3]; see also Bernheim and Rangel (2004, 2007a, 2007b, 2008).

As will be evident, my evaluation of neuroeconomics (as it pertains to standard ecoomics) is based in large part on the contemplation of research agendas that may or may not become technologically or practically feasible. My contention is only that there are conceivable paths to relevant and significant achievements, not that success is guaranteed. At this early stage in the evolution of neuroeconomics, the speculative visualization of such achievements is critical, both because it justifies the continuing interest and patience of mainstream economists, and because it helps neuroeconomists to hone more useful and relevant agendas.

A FRAMEWORK FOR DISCUSSION

While neuroeconomists are convinced that a better understanding of *how* decisions are made will lead to better predictions concerning *which* alternatives are chosen, many traditional economists greet that proposition with skepticism. Advocates and critics of neuroeconomics (as it pertains to standard economics) often appear to speak at cross-purposes, using similar language to discuss divergent matters, thereby rendering many exchanges largely unresponsive on both sides. In the earnest hope of avoiding such difficulties, I will first provide a framework for my discussion, so that I can articulate and address particular issues with precision.

Suppose our objective is to determine the causal effects of a set of environmental conditions, x, on a decision vector, y[4]. For the time being, we will take x to include only the types of variables normally considered by economists, such as income and taxes. We recognize nevertheless that y depends not only on x, but also on a set of unobservable conditions, ω, which may include variables of the type studied by neuroeconomists. We hypothesize that the causal relationship between y and the environmental conditions, (x, ω), is governed by some function f:

$$y = f(x, \omega) \tag{9.1}$$

It is important to emphasize that the function f could be either a simple reduced form (e.g., a demand

---

[1] See also Glimcher and Rustichini, 2004; Camerer *et al.*, 2005; Rustichini, 2005; Glimcher *et al.*, 2005; Rustichini, 2005; Camerer, 2007.

[2] The object of positive analysis is to make descriptive, testable statements concerning factual matters. In answering positive questions about decision making, an economist typically attempts to predict the alternatives an individual would choose under specified conditions.

[3] The objective of normative analysis is to make prescriptive statements – that is, statements concerning what *should* occur.

[4] Sometimes, the objective of traditional positive economics is simply to forecast y given a set of observed conditions x, without interpreting the forecasting relation as causal. In some contexts, it may be helpful to condition such forecasts on neuroeconomic variables; see for example the discussion on p. 120.

function expressing purchases of a good as a function of its own price, the prices of other goods, and income), or a more elaborate structural economic model. For instance, f could identify choices that maximize some objective function given the available alternatives when the conditions x and ω prevail[5].

Economists typically treat the unobserved conditions, ω, as noise, and attempt to determine the causal effects of the observed environmental conditions, x, on the *distribution* of decisions, y. If the distribution of ω is governed by a probability measure μ, then the distribution of y will correspond to a probability measure $\eta(\cdot\,|\,x)$, where for any Borel set A, $\eta(A\,|\,x) = \mu(\{\omega\,|\,f(x,\omega) \in A\})$. For example, the standard linear model assumes that

$$f(x,\omega) = x\beta + \varepsilon(\omega),$$

where ε is an unspecified function. It follow that $\eta(A\,|\,x) = \mu(\{\omega\,|\,x\beta + \varepsilon(\omega) \in A\})$.

Generally, economists attempt to estimate η directly from data on observable conditions, x, and decisions, y. In the case of the linear model, they estimate the parameter vector β along with parameters governing the distribution of $\varepsilon(\omega)$. There is no opportunity to recover the form of the function ε or the distribution of ω. Nor is there an obvious need. For example, when studying the behavioral effect of a sales tax on consumption, a traditional economist would not be concerned with quantifying the variation in consumption attributable to specific genetic traits; rather, he would focus on the distribution of responses (most notably the average) without conditioning on genetics. Accordingly, the identification of the causal relation $\eta(\cdot\,|\,x)$, where x consists of standard economic variables such as income and taxes, is arguably the primary objective of traditional positive economics.

In contrast, the objective of positive neuroeconomics is, in effect, to get inside the function f by studying brain processes. To illustrate, let's suppose that neural activity, z (a vector), depends on observed and unobserved environmental conditions, through some function Z:

$$z = Z(x,\omega)$$

Choices result from the interplay between cognitive activity the environmental conditions[6]:

$$y = Y(z,x,\omega)$$

It follows that

$$f(x,\omega) = Y(Z(x,\omega),x,\omega)$$

Positive neuroeconomics attempts to uncover the structure of the functions Z (the process that determines of neural activity) and Y (the neural process that determines decisions). Neuroeconomics necessarily treats the function f as a reduced form, even if it represents a structural economic model. Neuroeconomic research can also potentially shed light on the distribution of ω (the measure μ), which is the other component of η, the object of primary interest from the perspective of traditional positive economics.

The tasks of traditional positive economics and positive neuroeconomics are therefore plainly related. The question at hand is whether their interrelationships provide traditional positive economists with useful and significant opportunities to learn from neuroeconomics.

IS THE RELEVANCE OF NEUROECONOMICS SELF-EVIDENT?

Most members of the neuroeconomics community believe that the relevance of their field to economics is practically self-evident; consequently, they are puzzled by the persistent skepticism among mainstream economists. To motivate their agenda, they sometimes draw analogies to other subfields that have successfully opened "black boxes." For example, some liken neuroeconomics to the theory of the firm, which opened up the black box of production decisions (see Camerer *et al.*, 2004, 2005; Camerer, 2007).

From the perspective of a mainstream economist, analogies between neuroeconomics and the theory of the firm are misleading. In developing the theory of the firm, economists were not motivated by the desire to improve the measurement of reduced form production functions relating output to labor and capital. Rather, questions pertaining to the internal workings of the firm (unlike those pertaining to the internal workings of the mind) fall squarely within the historical boundaries of mainstream economics, because they

---

[5] In the latter case, an economist would typically interpret the free parameters of the objective function as aspects of preferences. However, modern choice theory teaches us that preferences and utility functions are merely constructs that economists invent to summarize systematic behavioral patterns. We are of course concerned with the accurate estimation of those parameters, but only because they allow us to recover the behavioral relation f.

[6] The arguments of Y include x and ω in addition to z because the same neural activity could lead to different outcomes depending on the environmental conditions.

concern the nature of organized exchange between individuals. An economist who seeks to understand prices, wages, risk-sharing, and other traditional aspects of resource allocation has an undeniable stake in understanding how trade plays out within a range of institutions, including markets and firms, and how different types of exchange come to be governed by different types of institutions. In contrast, the mind is not an economic institution, and exchange between individuals does not take place within it.

Notably, economists have not materially benefited from a long-standing ability to open up other black boxes. For example, we could have spent the last hundred years developing highly nuanced theories of production processes through the study of physics and engineering, but did not. A skeptical mainstream economist might also note that models of neural processes are also black boxes. Indeed, the black-box analogy is itself false: we are dealing not with a single black box, but rather with a Russian doll. Do we truly believe that good economics requires a command of string theory?

It is therefore understandable that so many economists are unmoved by the amorphous possibility that delving into the nuts and bolts of decision making will lead to better and more useful economic theories. To persuade them that a particular black box merits its opening, one must at least provide a speculative roadmap, outlining reasonably specific potentialities which economists would recognize as both exciting and within the realm of possibility. What has been offered along these lines to date is far too vague and insubstantial to convert the skeptics.

SOME SPECIFIC SOURCES OF SKEPTICISM

Neuroeconomists have certainly attempted to offer economists a variety of affirmative motivations for opening the black box of the human mind. Many mainstream economists find those motivations unpersuasive because they see neuroeconomic inquiry as largely orthogonal to traditional economic analysis, a view that finds its most forceful articulation in the work of Gul and Pesendorfer (2008). To identify motivations that economists would generally find persuasive, one must first understand the logic of that view, and appreciate its appeal.

Much of the prevailing skepticism concerning the magnitude of the contribution that neuroeconomics can potentially make to standard positive economics arises from the following three considerations.

First, unless neuroeconomics helps us recover the behavioral relation η, its contributions will not advance the historical objectives of positive economics. Though the functions Y and Z are obviously interesting, the questions they address directly are not ones that mainstream economists traditionally examine.

Second, because the behavioral relation η involves no neural variables, traditional positive economists can divine its properties from standard economic data. Distinguishing between two neural processes, (Y, Z, μ) and (Y', Z', μ'), is helpful to such an economist only if the differences between those processes lead to significant differences between the corresponding reduced form representations, η and η'. But if the latter differences are indeed significant, then an economist can test between η and η' directly using standard economic data, without relying on neuroeconomic methods.

Third, while neuroeconomics potentially offers another route to uncovering the structure of the relation η, there is skepticism concerning the likelihood that it will actually improve upon traditional methods. The prospects for building up a *complete* model of complex economic decisions from neural foundations would appear remote at this time. Even if such a model were assembled, it might not be especially useful. Precise algorithmic models of decision making of the sort to which many neuroeconomists aspire would presumably map highly detailed descriptions of environmental and neurobiological conditions into choices. In constructing the distribution η from Y, Z, and μ, a microeconomist would treat vast amounts of this "micro-micro" information as noise. An economist might reasonably hope to apprehend the structure of η more readily by studying the relationship between y and x directly, particularly if the explanatory variables of interest (x) include a relatively small number of standard environmental conditions. As an example, suppose η is the household demand function for a good. What does a standard economist lose by subsuming all of the indiosyncratic, micro-micro factors that influence decisions, many of which change from moment to moment, within a statistical disturbance term? What can neuroeconomics teach us about the relationship between average purchases and the standard economic variables of interest (prices, income, and advertising) that we cannot discern by studying those relationships directly?

These considerations do not, however, rule out the possibility that neuroeconomics might make significant contributions to mainstream economics. With respect to the second consideration, even the most skeptical economist must acknowledge that the standard data required to address questions of interest are sometimes unavailable, and are rarely generated under ideal conditions. Surely we should explore the possibility that new types of data and methods of analysis might help us overcome those limitations. Thus, the third consideration emerges as the most central to my appraisal, and the rest of this chapter is devoted to its evaluation.

In principle, even without providing a complete neural model of complex economic decision making, neuroeconomics offers several potential routes to uncovering the structure of standard behavioral relationships. First, it will lead to the measurement of new variables, which may usefully find their way into otherwise standard economic analyses. I discuss that possibility in the next two sections. Second, detailed knowledge concerning the neural processes of decision making may help economists discriminate between theories and/or choose between models. As discussed on pp. 120–121, the formulation of rigorous tests may prove challenging. Standard economic theories of decision making concern choice patterns, and are therefore agnostic with respect to decision processes; hence, they may have few testable neural implications. The penultimate two sections examine the more modest possibility that understanding a neural process may provide economists with informal but nevertheless useful guidance with respect to model selection (specifically, explanatory variables and functional forms).

A skeptic might observe that the most promising routes to meaningful contributions are also the most limited. An economist who examines neural variables would not necessarily require extensive knowledge of neuroeconomic methods or a deep appreciation of neural processes; instead, he might simply rely on neuroeconomists to identify and collect the relevant data. Similarly, even if findings from neuroscience informally guide aspects of model selection (variables and/or functional forms), once a traditional positive economist knows the structure of the selected model, he can discard all information concerning neural processes without loss.

Many psychologists would view the positions outlined above as a form of radical behaviorism. They are surprised that economists still hew so rigidly to a perspective that psychology abandoned decades ago. Yet the different paths of psychology and economics are not so difficult to understand once we consider divergent objectives of those disciplines. I would point to two important differences. First, unlike economics, the field of psychology has traditionally subsumed questions about the mind. Thus, traditional psychological questions pertain to aspects of the functions Y and Z, whereas traditional economic questions do not. Second, questions in psychology often focus on the micro-micro determinants of behavior. A psychologist is potentially interested the particular factors that cause a single individual to behave in a certain way at a specific moment. In contrast, traditional economic analysis usually treats such idiosyncratic influences as background noise.

ARE THERE USES FOR EXOGENOUS NEUROECONOMIC VARIABLES?

The discussion above takes $\eta(\cdot\,|\,x)$, with x defined to include only traditional economic variables, as the object of interest for traditional positive economics. It therefore ignores the possibility that neuroeconomics might redraw the boundary between the set of variables that economists treat as observable (x), and those they treat as unobservable (ω). More formally, by measuring some vector of variables $\tilde{\omega}$, a neuroeconomist can repartition the environmental conditions (x, ω) into (x^0, ω^0), where $x^0 = (x, \tilde{\omega})$ and $\omega = (\omega^0, \tilde{\omega})$, and potentially allow economists to recover the causal relation $\eta^0(\cdot\,|\,x^0)$. It is important to acknowledge that the barriers to redrawing this boundary may be practical and political (e.g., privacy concerns), not merely technological. For the purpose of this discussion, let us suspend disbelief and consider the possibilities.

Why might the distribution $\eta^0(\cdot\,|\,x^0)$, which subsumes the behavioral effects of neural variables, as well as the effects of standard environmental factors conditional on neural variables, be of interest to mainstream economists? The answer is not obvious. Suppose a neuroeconomist discovers a genetic trait that helps predict saving (a "patience gene"). Should economists greet that discovery with enthusiasm? Economics has not concerned itself historically with the relationship between genetics and saving. An economist might question whether that knowledge is likely to improve his understanding of the effects of, say, capital income taxes (an element of x) on asset accumulation, averaged or aggregated over the elements of ω (including genetics).

Further reflection suggests, however, that exogenous neural variables may have a variety of uses. For a more complete discussion of possible uses, along with examples, see Bernheim (2008). First, neural proxies for tastes and talents may facilitate the detection of biases arising from omitted variables, and the inclusion of such proxies in econometric specifications may mitigate omitted variables bias. Second, when the decisions of several distinct individuals are co-determined (as in peer groups), we may be able to measure the causal effect of one individual's choice on another's decision by using the first individual's exogenous neural predispositions as instruments. Third, if an economist is narrowly concerned with forecasting behavior as of a particular moment in time, and if a time-varying neural condition is known to affect the behavior in question, then the use of information concerning that condition can improve the forecast. Fourth, causal relationships that are conditioned on neural characteristics may be useful when projecting the effects of a policy from one population to another, particularly if the two populations differ compositionally. Fifth, understanding the roles of genetic predispositions in decision making may shed light on the likely sensitivity of behavior to policy interventions. Sixth, if private firms begin to measure the neural characteristics of consumers or employees and use that information in the course business, economists will need to consider the roles of neural variables in resource allocation. Even if governments prevent such activities due to privacy concerns, economists will be unequipped to evaluate the effects of such policies unless they study the neural correlates of behavior.

ARE THERE USES FOR ENDOGENOUS NEUROECONOMIC VARIABLES?

As I explained earlier, one of the main objectives of neuroeconomics is to uncover the structure of the function Y, which maps endogenous neural activity, z, along with the environmental conditions x and ω, to decisions. Based on existing findings concerning Y, it is already possible to predict certain choices from particular types of endogenous neural activity with a high degree of accuracy. For examples, see Knutson *et al.* (2007), Kuhnen and Knutson, (2005), and Hsu *et al.* (2005). Because accurate behavioral prediction is a central goal of positive economics, many neuroeconomists have offered such findings as evidence of their field's relevance (see, for example, Camerer, 2007).

Why are mainstream economists unpersuaded by this evidence? In the context of most traditional economic questions, they see little value in predicting behavior based on its endogenous components (here, z). Consider the following stark example. Suppose our goal is to predict whether individual customers at a grocery store will purchase milk. After carefully studying a large sample of customers, a confused graduate student declares success, noting that it is possible to predict milk purchases accurately with a single variable: whether the customer reaches out to grab a carton of milk. The technology to collect this highly predictive data has long been available; economists have demurred not due to a lack of creativity, boldness, and vision, but rather because such predictions are of no value to them.

Mainstream economists should not, however, completely dismiss the possibility that endogenous neural variables will prove useful. In some situations, information concerning some aspect of the environmental conditions, x, or the decision, y, may not be available. Data on neural activity (z) along with knowledge of the functions Y and Z can then potentially permit us to impute the missing conventional variables, and use the imputed values in otherwise standard economic analyses. For example, the analysis of Wang *et al.* (2006) suggests that it may be possible to infer private information concerning standard economic variables from neural responses. See Bernheim (2008) for a more detailed discussion of the possibilities for imputing both exogenous variables and choices.

DO ECONOMIC THEORIES HAVE TESTABLE IMPLICATIONS CONCERNING NEURAL PROCESSES?

Perhaps the most tantalizing claim concerning the potential prospects of neuroeconomics is that an understanding of neural processes may provide economists with new opportunities to formulate direct tests of both standard and nonstandard (behavioral) theories of decision making (see, e.g., Camerer, 2007)[7]. While such advances are conceivable, it is important for neuroeconomists to acknowledge the difficulty of that endeavor, and to avoid premature conceptual leaps, especially if they hope to be taken seriously by mainstream economists.

The central conceptual difficulty arises from the fact that standard economic theory (including neoclassical economics as well as much of modern behavioral economics) is agnostic with respect to the nature of decision *processes*. No explicit assumptions are made concerning the inner workings of the brain. For example, contrary to the apparent belief of many non-economists, economists do not proceed from the premise that an individual literally assigns utility values to alternatives, and from any opportunity set chooses the alternative with the highest assigned value. This disciplinary agnosticism with respect to process accounts for Gul and Pesendorfer's (2008) contention that neural evidence cannot shed light on standard economic hypotheses.

---

[7] This issue is distinct from the possibility that the measurement of neural variables may facilitate tests of conventional economic theories, e.g. by providing instruments or permitting reliable imputations for missing variables. The question here is whether one can test an economic theory of behavior by examining the *process* that governs decision making.

Foundational economic assumptions concern choice patterns, not processes. Neoclassical decision theory follows from a collection of choice axioms, the most critical of which is sometimes labeled *independence of irrelevant alternatives* (a generalization of the more familiar weak axiom of revealed preference). According to that axiom, if an individual chooses a particular alternative from an opportunity set, then he will also choose that alternative from any smaller set, provided the alternative remains available. When the independence axiom is satisfied, there exists an ordering (interpreted as preferences) that rationalizes all of the individual's choices, in the sense that he always chooses the most highly ranked alternative according to the ordering. With some additional (but largely technical) axioms, one can also represent his choices as maximizing a continuous utility function. Within this framework, preferences and utility are merely constructs, invented by the economist to provide a convenient representation of choice patterns. The theory does not assert that these constructs have counterparts within the brain. Consequently, those who would test the theory by searching for such counterparts have misunderstood the theory's foundations.

The preceding observations do not, however, imply that neural evidence is conceptually incapable of shedding light on standard economic hypotheses. Choice axioms cannot be valid unless the neural processes that govern choice are capable of delivering decisions that conform to the axioms; thus, a mainstream economist cannot remain *entirely* agnostic as to process. To take an extreme possibility, if neuroeconomists succeed in reducing all pertinent neural decision processes to a precise computational algorithm for some reasonably large class of decision problems, they will be able to determine whether the algorithm delivers choices that satisfy the independence axiom, and thereby test neoclassical decision theory. However, that potentiality does not convincingly establish the value of neuroeconomics, for two reasons.

First, assume we have reason to believe that the brain sometimes employs a particular decision algorithm, but have not yet established the scope of that algorithm's application. Suppose the algorithm's implications for choice within some domain of decision problems, A, would be inconsistent with some economic theory; moreover, there is no subset of A for which the same statement holds. We might hope to disprove the economic theory by demonstrating that the decision algorithm in fact governs choices throughout the domain A. However, a formal test of the latter hypothesis would presumably involve a comparison between the algorithm's behavioral predictions and actual choices throughout A. But if data on those decisions are available, we can test the economic theory directly, without concerning ourselves with the nuts and bolts of decision processes. Thus, the incremental contribution of neuroeconomics is not obvious.

Second, neuroeconomics is still a long way from reducing the neural processes that govern the complex decisions with which economists are conventionally concerned to precise algorithms, especially for broad classes of environments. Existing algorithmic representations of such processes pertain only to very simple tasks and functions. Much of what is known has a qualitative flavor – for example, that certain types of decisions involve elevated activity in particular regions of the brain, and that those regions tend to be associated with specific functions. While it is conceivable that we might be able to test economic theories using such information, the necessary conceptual groundwork for such a test has not yet been laid. See Bernheim (2008) for a more detailed dicsussion, including an analysis of what that groundwork would entail.

Unfortunately, the neuroeconomic community has not yet generally acknowledged the conceptual challenges that one necessarily confronts when attempting to derive testable implications of economic theories for neural processes. Instead, neuroeconomists have sometimes proceeded (at times implicitly) as if those implications are obvious or easily motivated. That practice leaves many mainstream economists with the regrettable (and often inaccurate) impression that neuroeconomists do not adequately understand the economic theories upon which they hope to shed light. I discuss three examples in Bernheim (2008): the contention that McClure *et al.* (2004) provided a neural test of quasihyperbolic discounting, the claim that Harbaugh *et al.* (2007; see also Chapter 20 of this volume) tested theories of altruism and "warm glow" giving, and the notion that the evidence in Platt and Glimcher (1999) supports expected utility theory. As I explain, none of those claims withstands scrutiny.

CAN AN UNDERSTANDING OF NEURAL PROCESSES USEFULLY GUIDE MODEL SELECTION?

The number of empirical models that an economist could construct to describe any particular decision as a function of conventional explanatory variables is vast. Even if neuroeconomics does not provide new variables of interest or an independent foundation for testing one model against another, it could conceivably generate suggestive findings that informally guide the search for an appropriate empirical model in useful directions, leading to more rapid and effective identification of the best predictive relationship. Here, the two main aspects of model selection are discussed: variable selection and the choice of functional form.

First, consider variable selection. Neuroeconomic evidence could in principle motivate the inclusion of particular conventional variables in specific behavioral models. Suppose, for example, that mandated transfers to others influence brain activity in centers linked to reward-processing (Harbaugh *et al.*, 2007; see also Chapter 20 of this volume). While such evidence would not prove that altruism motivates behavior, it might well *suggest* such a hypothesis to an empirical economist, who might then investigate behavioral models that incorporate related variables (e.g., measures of potential external effects). Similarly, an examination of neural evidence concerning the processes that govern attention might suggest that consumers are potentially susceptible to tax illusion, and that they will respond differently depending on whether a product is tagged with tax-inclusive or tax-exclusive prices. Such evidence might lead an empirical economist to examine empirical models that separately include explanatory variables measuring posted prices and hidden taxes.

While acknowledging the possibility described in the preceding paragraph, a skeptic might nevertheless question whether neuroeconomics is likely to make such contributions in practice. Empirical economists have other sources of guidance and inspiration, such as introspection and research from psychology. Indeed, neural studies such as that by Harbaugh *et al.* (2007; see also Chapter 20) are themselves motivated by hypotheses imported from other fields. Likewise, economists formulated and tested conjectures concerning tax illusion based on a common-sense understanding of attention, without the benefit of neuroeconomic evidence; see in particular Chetty et al. (2007), and Finkelstein (2007). Empirical economists who are not persuaded to investigate the roles of pertinent variables in behavioral relationships on the basis of other considerations are unlikely to find neural evidence convincing. To uniquely motivate the inclusion of a potential explanatory variable that empirical economists have ignored, a neuroeconomist would literally have to stumble across some unexpected environmental correlate of brain activity. I do not dismiss that possibility, but neither does it convince me that the field holds great potential for conventional positive economics.

Even if research on the neurobiology of decision making had provided the impetus for investigating altruism, tax illusion, or some other phenomenon, it seems unlikely that an empirical strategy for estimating the function η would have been influenced by the details of the neurobiological evidence. Rather, that evidence would have merely *motivated* (to use Gul and Pesendorfer's term) an examination of functional forms that include the pertinent variables. It is not at all obvious that an economist who possesses a deep understanding of the motivating scientific evidence would be any better equipped to estimate η than one who simply apprehends the pertinent psychological principles intuitively.

In addition to suggesting that certain variables may play roles in particular behavioral relationships, neuroeconomic evidence may also indicate that others play no role. Such evidence could motivate exclusion restrictions. Indeed, formal neural tests of exclusion restrictions are conceivable in principle, even without precise knowledge of the computational algorithms that govern decision making. We can frame the issue as a computer-programming task. To implement a choice mapping that depends on a particular variable, computer code must reference that variable. For any neural process that implements the same computational algorithm, there must presumably be some neural response to the variable's value. Consequently, the absence of any response would formally justify an exclusion restriction in the behavioral relationship.

Next, consider the choice of functional form. In principle, the nature of neurobiological response mechanisms may suggest particular empirical specifications. For example, there is some evidence that temporal difference reinforcement learning (TDRL) models accurately describe the operation of neural systems governing dopamine learning (Schultz *et al.*, 1997; Schultz, 1998, 2000). These parsimonious, tightly parameterized learning models could guide the formulation of empirical behavioral relationships in settings that involve the accumulation of experience. Because other learning processes may also influence choices, the neural evidence cannot *prove* that one functional form is better than another for the purpose of predicting behavior. However, it could lead economists to examine particular parsimonious specifications that they might not otherwise consider, and some of these may outperform more conventional alternatives.

A mere catalog of such possibilities will never suffice to convince the skeptics, nor should it. Mainstream economists should acknowledge the conceptual possibilities discussed above, and exercise intellectual tolerance and patience while neuroeconomists explore them. Neuroeconomists in turn should recognize that the burden of proof is squarely on their shoulders. Skeptical reactions define a specific challenge: *Provide* *an example of a novel economic model derived originally from neuroeconomic research that improves our measurement of the causal relationship between a standard exogenous environmental condition – one with which economists have been historically concerned – and a standard economic choice*. Unless the neuroeconomics community eventually rises to that challenge, the possibilities discussed in this section will eventually be dismissed as unfounded speculation.

CAN NEUROECONOMICS IMPROVE OUT-OF-SAMPLE PREDICTIONS?

Sometimes, economists wish to predict behavior under completely novel conditions (for example, a new and untried public policy). There is no assurance that reduced form behavioral models will perform well in such contexts, especially if the novel conditions are qualitatively distinct from any that have preceded them. In contrast, a good structural model, based on a deeper understanding of behavior, may permit reasonable projections even when fundamental environmental changes occur. Many neuroeconomists believe that their field will provide such models.

By way of analogy, suppose a computer has been programmed to make selections for choice problems falling into a number of distinct categories, but the tasks for which we have observed its choices belong to a subset of those categories. We could potentially develop a good positive model that predicts the computer's choices for problems within the categories for which we have data. However, based on that limited data, projecting choices for problems within the remaining categories is guesswork. Now suppose we obtain the computer code. In that case, even without additional choice data, we could accurately predict the computer's decisions in *all* circumstances. When neuroeconomists suggest that an understanding of the brain's computational algorithms will permit more reliable out-of-sample behavioral predictions, they are making an analogous claim.

Unfortunately, the issue is not quite so straightforward. If neuroeconomists only succeed in mapping a subset of the brain's neural circuitry to computational algorithms, out-of-sample prediction will remain problematic. To pursue the analogy to a computer program a bit further, suppose we obtain the code only for certain subroutines that are activated when the computer solves problems falling within the categories for which we have data. There is no guarantee that it will activate the same subroutines for related purposes when confronting problems within the remaining categories, particularly if those problems are qualitatively different from the ones previously encountered. Without knowing how the entire program operates, including the full array of subroutines upon which it can call, as well as the conditions under which it activates each of them, one cannot simulate its operation in fundamentally new environments.

Of course, one can proceed based on the *assumption* that the brain will continue to use the same neural circuitry in the same way when confronting new classes of decision problems. But there is no guarantee of greater out-of-sample stability at the neural level than at the behavioral level[8]. Whether we would be better off making out-of-sample predictions from structural neural models rather than structural behavioral models is therefore a factual question that can only be settled through experience, and not through logical arguments.

Still, there are reasons to hope that consideration of evidence on neural processes might at least help us select *economic* models that are more reliable for the purpose of making out-of-sample projections. Imagine, for example, that an estimated within-sample behavioral relationship is equally consistent with several distinct structural economic models, each of which has a different out-of-sample behavioral implication. Suppose the available neural evidence informally persuades us (but does not prove) that one of those models is more likely to match reality. Then we might reasonably hope to obtain more accurate out-of-sample predictions from the preferred model.

Consider the following example. Currently, tens of millions of people lack health insurance coverage. One theory holds that those households have carefully assessed the costs and benefits of insurance, and concluded that it is too costly; another holds that they are inattentive to their health-care needs, and hence unresponsive to costs and benefits. Both hypotheses are equally consistent with observed choices, but they have starkly different out-of-sample implications concerning the fraction who would purchase insurance if the cost of coverage were reduced well below historical levels. Can neuroeconomics help us judge between their divergent predictions? Suppose we use

---

[8] Just as a structural economic model can be viewed as a reduced form for a structural neural model, any structural neural model can also be viewed as a reduced form for some deeper structure, and the stability of the neural reduced form over classes of environments will depend on how that deeper structure operates. If, for example, secondary neural systems are designed to override a primary system whenever the latter would generate behavior too far from some norm, then an incomplete neural model of choice might be less stable out of sample than a behavioral model.

neural methods to measure attentiveness to health-care needs, as well as value assessments for insurance coverage. The first theory informally predicts high attentiveness and high-value assessments; the second has the opposite prediction. Neither finding would *prove* that the uninsured are more likely to behave one way or the other out of sample. For example, the uninsured might *start* attending to health-care issues and contemplating the benefits of insurance if they thought health care was affordable. Even so, the neural evidence would presumably influence our comfort with and degree of confidence in each model.

These possibilities are of course speculative. Mainstream economists will relinquish their skepticism only when confronted with examples of superior out-of-sample prediction in contexts involving the types of environmental conditions and behaviors that they ordinarily study.

CONCLUSIONS

The potential for the emerging field of neuroeconomics to shed light on traditional economic questions has been overstated by some, unappreciated by others, and misunderstood by many. With respect to positive economics, the case for studying the neural foundations of decision making is hardly self-evident. Nevertheless, neuroeconomics could in principle contribute to conventional positive economics in a number of ways, which I have attempted to catalog.

At the same time, a number of the potential contributions discussed in this paper strike me as somewhat modest, rather special, and/or somewhat peripheral. While there is good reason to hope that some of the contributions will prove noteworthy, I have considerably more difficulty convincing myself that neuroeconomics is likely to become a central or indispensable component of standard positive economics, or that it will revolutionize the field in some fundamental way. Whether that assessment reflects the field's actual limitations or the deficient imagination of a relatively mild skeptic remains to be seen.

Due to space constraints, I have not evaluated potential contributions to normative economics. I doubt that neuroeconomics will provide a technology for measuring utility directly, or that it will replace choice-based welfare analysis with a new utilitarian paradigm. However, it may hold the potential to improve choice-based welfare analysis; see Bernheim (2008) for a detailed discussion.

Many neuroeconomists have been surprised and frustrated to learn that skepticism concerning their field's potential among mainstream economists runs deep. How can they combat that skepticism? First, neuroeconomists need to do a better job of articulating specific visions of the field's potential contributions to mainstream economics. Such an articulation would ideally identify a standard economic question of broad interest (e.g., how taxes affect saving), and outline a conceivable research agenda that could lead to specific, useful insights of direct relevance to that question. Vague assertions that a deeper understanding of decision-making processes will lead to better models of choice do not suffice. Second, it is essential to avoid hyperbole. Exaggerated claims simply fuel skepticism. Sober appraisals of the field's potential, including its limitations, will promote its acceptance more effectively than aggressive speculation that involves loose reasoning or otherwise strains credibility. Third, the ultimate proof is in the pudding. To convert the skeptics, neuroeconomists need to accumulate the right type of success stories – ones that illuminate conventional economic questions that attracted wide interest among economists prior to the advent of neuroeconomic research.

Acknowledgments

This paper was prepared for a symposium titled "Neuroeconomics: Decision Making and the Brain," held at NYU on January 11–13, 2008. I am grateful to Antonio Rangel and Colin Camerer for stimulating discussions and comments. I also acknowledge financial support from the National Science Foundation through grant number SES-0452300.

References

Bernheim, B.D. (2008). *Neuroeconomics: a sober (but hopeful) appraisal*. AEJ: Microeconomics, in press.

Bernheim, B.D. and Rangel, A. (2004). Addiction and cue-triggered decision processes. *Am. Econ. Rev.* 94, 1558–1590.

Bernheim, B.D. and Rangel, A. (2007a). *Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics*. Mimeo, Stanford University.

Bernheim, B.D. and Rangel, A. (2007b). Toward choice-theoretic foundations for behavioral welfare economics. *Am. Econ. Rev. Papers Proc.* 97, 464–470.

Bernheim, B.D. and Rangel, A. (2008). Choice-theoretic foundations for behavioral welfare economics. In: A. Caplin and A. Schotter (eds), *The Methodologies of Modern Economics*. Oxford: Oxford University Press in press.

Camerer, C.F. (2007). Neuroeconomics: using neuroscience to make economic predictions. *Economic J.* 117, C26–C42.

Camerer, C.F., Loewenstein, G., and Prelec, D. (2004). Neuroeconomics: why economics needs brains. *Scand. J. Economics* 106, 555–579.

Camerer, C.F., Loewenstein, G., and Prelec, D. (2005). Neuroeconomics: how neuroscience can inform economics. *J. Econ. Lit.* 43, 9–64.

Chetty, R., Looney, A., and Kroft, K. (2007). *Salience and taxation: theory and evidence*. Mimeo, University of California, Berkeley.

Finkelstein, A. (2007). *EZ-tax: tax salience and tax rates*. Mimeo, MIT.

Glimcher, P.W., Dorris, M.C., and Bayer, H.M. (2005). Physiological utility theory and the neuroeconomics of choice. *Games Econ. Behav.* 52, 213–256.

Glimcher, P.W. and Rustichini, A. (2004). Neuroeconomics: the consilience of brain and decision. *Science* 306, 447–452.

Gul, F. and Pesendorfer, W. (2008). In: A. Caplin and A. Schotter (eds.), *The Methodologies of Modern Economics*. Oxford, Oxford University Press. In press.

Harbaugh, W.T., Mayr, U., and Burghart, D.R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* 316, 1622–1625.

Hsu, M.M., Bhatt, M., Adolphs, R. *et al*. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310, 1680–1683.

Knutson, B., Rick, S., Elliott Wimmer, G. *et al*. (2007). Neural predictors of purchases. *Neuron* 53, 147–156.

Kuhnen, C.M. and Knutson, B. (2005). The neural basis of financial risk taking. *Neuron* 47, 763–770.

McClure, S.M., Laibson, D.I., Loewenstein, G., and Cohen, J.D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507.

Platt, M.L. and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.

Rustichini, A. (2005). Neuroeconomics: present and future. *Games Econ. Behav.* 52, 201–212.

Savage, L. (1954). *The Foundation of Statistics*. New York, NY: John Wiley & Sons.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.

Schultz, W. (2000). Multiple reward signals in the brain. *Nat. Rev. Neurosci.* 1, 199–207.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Wang, J.T.-Y., Spezio, M., and Camerer, C.F. (2006). *Pinocchio's pupil: using eyetracking and pupil dilation to understand truth-telling and deception in biased transmission games*. Mimeo, Caltech.

CHAPTER

10

Decisions Under Uncertainty: Psychological, Economic, and Neuroeconomic Explanations of Risk Preference

Elke U. Weber and Eric J. Johnson

OUTLINE

RISK PREFERENCE: THE HISTORICAL CONTEXT

Democratic and libertarian societies ask their citizens to make many decisions that involve uncertainty and risk, including important choices about pension investments, medical and long-term care insurance, or medical treatments. Risky decisions, from barely conscious ones when driving ("Should I overtake this car?") to carefully deliberated ones about capital investments ("Do I need to adjust my portfolio weights?") abound. As citizens have taken on more decision responsibility, unpredictability and uncertainty of decision outcomes has increased as the result of ever faster social, institutional, environmental and technological change.

It is no surprise, then, that the topic of decision making under risk and uncertainty has fascinated observers of human behavior. From philosophers charged with providing tactical gambling advice to noblemen, to economists charged with predicting people's reactions to tax changes, risky choice and the selection criterion that people seek to optimize when making such decisions has been the object of theoretical and empirical investigation for centuries (Machina, 1987; Glimcher, 2003).

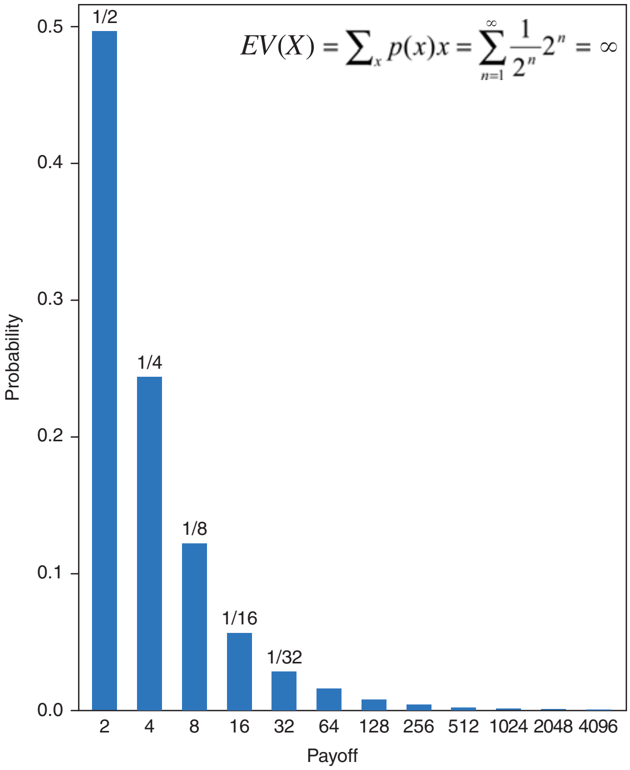


FIGURE 10.1 Payoff distribution for St Petersburg paradox game, where a fair coin is tossed until the first "head" is scored. The payoff depends on the trial at which the first "head" occurs, with \$2 if on the first trial, \$4 if on the second trial, and \$$2^n$ if on the nth trial.

Expected Value Theory

The maximization of expected (monetary) value (EV) of gamble X,

$$EV(X) = \sum_x p(x) \cdot x, \tag{10.1}$$

first considered in the mid-seventeenth century, was rejected as a universally applicable decision criterion based on the so-called St Petersburg paradox, where people are willing to pay only a small price (typically between \$2 and \$4) for the privilege of playing a game with a highly skewed payoff distribution that has infinite expected value, as shown in Figure 10.1.

Expected Utility Theory

To resolve the St Petersburg paradox, Bernoulli (1954/1738) proposed that people maximize expected utility (EU) rather than expected value,

$$EU(X) = \sum_x p(x)u(x), \tag{10.2}$$

postulating that money and wealth are diminishing in value, as shown in Figure 10.2. The function that maps actual wealth (x) on the x-axis into utility for wealth ($u(x)$) is no longer linear but "concave." An increase in wealth of \$1000 is worth a lot more at lower initial levels of wealth (from \$0 to \$1000) than at higher initial levels (from \$2000 to \$3000). In power functions, $u(x) = x^{\theta}$, for example, the exponent θ is a parameter that describes the function's degree of curvature ($\theta = .50$ in Figure 10.2) and serves as an index of an individual's degree of risk aversion. Such an individual difference parameter has face validity, as some individuals seem to resolve choices among options that differ in risk in very cautious ways ($\theta < 1$), while others seem willing to take on great risks in the hope of even greater returns ($\theta > 1$).

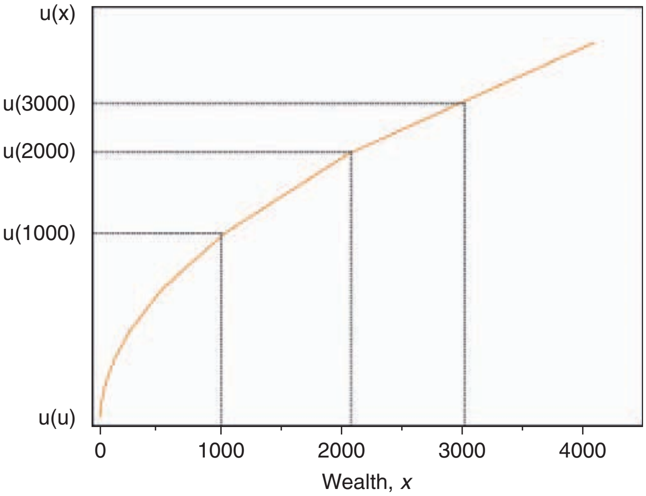


FIGURE 10.2 Concave utility function $u(x) = x^{0.5}$ which converts wealth, x, into its utility $u(x)$. An increase in wealth from \$0 to \$1000 is shown to result in a greater increase in utility than an increase in wealth from \$2000 to \$3000.

Von Neumann and Morgenstern (1947) provided an intuitively appealing axiomatic foundation for expected utility (EU) maximization, which made it a normatively attractive decision criterion not only for repeated decisions in the long run but also for unique risky decisions, and made EU maximization the dominant assumption in the economic analysis of choice under risk and uncertainty. See Chapter 3 of this volume for more detail on EU theory and its variants.

Risk–Return Models

In parallel to these developments in economics, Markowitz (1959) proposed a somewhat different solution to the St Petersburg paradox in finance,

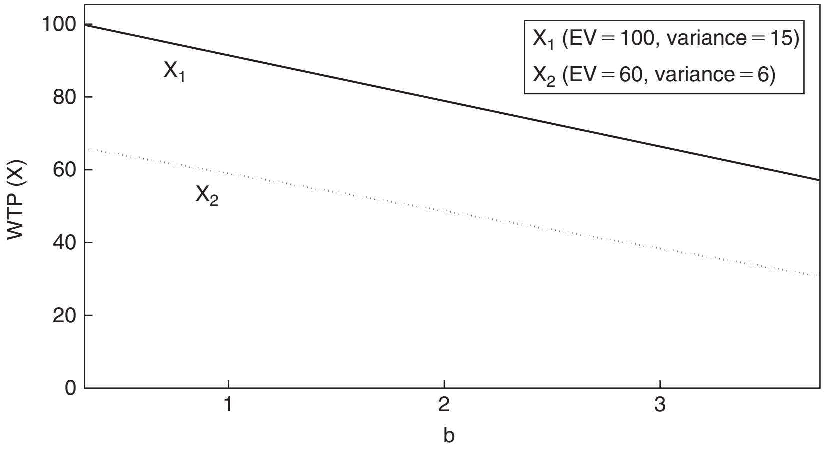


FIGURE 10.3 Willingess-to-pay (WTP) for risky investment options X (for X_1 ($EV = 100$, Variance = 15) and X_2 ($EV = 60$, Variance = 6)) as predicted by risk-return model in Equation 10.3, for different values of b.

modeling people's willingness to pay (WTP) for risky option X as a tradeoff between the option's return $V(X)$ and its risk $R(X)$, with the assumption that people will try to minimize level of risk for a given level of return:

$$\text{WTP}(X) = \text{V}(X) - b\text{R}(X). \tag{10.3}$$

Traditional risk–return models in finance equate $V(X)$ with the EV of option X and $R(X)$ with its variance. Model parameter b describes the precise nature of the tradeoff between the maximization of return and minimization of risk, and serves as an individual difference index of risk aversion. Figure 10.3 shows how WTP varies for two gambles as a function of the tradeoff parameter b. This risk–return tradeoff model is widely used in finance, e.g., in the Capital Asset Pricing Model (CAPM; Sharpe, 1964; see Bodie and Merton, 1999, for more detail), and can be seen as a quadratic approximation to a power or exponential utility function (Levy and Markowitz, 1979). Other classes of utility functions also have risk–return interpretations, where returns, V(X), are typically modeled as the EV of the risky option, and different utility functions imply different functional forms for risk, $R(X)$ (Jia and Dyer, 1997).

Despite their prescriptive and normative strengths, both EU maximization and risk–return optimization have encountered problems as descriptive models for decisions under risk and uncertainty. Experimental evidence as well as choice patterns observed in the real world suggests that individuals often do not behave in a manner consistent with either of these classes of models (McFadden, 1999; Camerer, 2000). Human choice behavior deviates in systematic ways, as captured originally in two classical demonstrations referred to as the Allais (1953) and Ellsberg (1961) paradoxes, described below.

Limitations of Economic Risky Choice Models

A central assumption of all economic risky choice models described above is that the utility of decision outcomes or the risk and return of choice options are determined entirely by the objective value of possible outcomes (and the final wealth they generate) in a "reference-independent" way, i.e., in a way that does not depend on what the outcome can be compared to. Thus the receipt of a $100 is assumed to have the same effect on the decision of an individual, whether is it the top prize in the office basketball pool or the consolation prize in a lottery for one million dollars. Decision-makers' evaluation of outcomes and choice options, however, appears to be influenced by a variety of relative comparisons (Kahneman, 2003).

In fact it is now widely known that people often compare the outcome of their chosen option with the outcome they could have gotten, had they selected a different option (Landman, 1993). Such comparisons have an obvious learning function, particularly when the "counterfactual" outcome (i.e., the outcome that could have been obtained, but wasn't) would have been better. This unfavorable comparison between what was received and what could have been received with a different (counterfactual) action under the same state of the world is termed *regret*. When the realized outcome is better than the alternative, the feeling is called *rejoicing*. Consistent with the negativity effect found in many judgment domains (Weber, 1994), feelings of regret are typically stronger than feelings of rejoicing. Regret theory, independently proposed by Loomes and Sugden (1982) and Bell (1982), assumes that decision makers anticipate these feelings of regret and rejoicing, and attempt to maximize EU as well as minimizing anticipated post-decisional net regret. Susceptibility to regret is a model parameter and an individual difference variable that dictates the specifics of the tradeoff between the two choice criteria. Minimization of anticipated decision regret is a goal frequently observed, even if it results in lower material profitability (Markman *et al.*, 1993). Extending these ideas, Braun and Muermann (2004) proposed a formulation of regret theory that can be applied to decisions that have more than two possible choice options. While post-decisional regret undoubtedly plays an important learning function, the importance to pre-decisional, anticipated regret is less clear. A recent set of choice simulations by Laciana *et al.* (2007) showed that the incorporation of anticipated regret into *EU* maximization did not result in risky choices that were significantly differently from those of *EU* maximization in a real-world risky decision domain, namely precision agriculture. In contrast, the actions prescribed by prospect theory value maximization, a theory described next, were considerably different from those prescribed by *EU* maximization.

Prospect Theory

Prospect theory (PT; Kahneman and Tversky, 1979; Tversky and Kahneman, 1992) introduced a different type of relative comparison into the evaluation of risky choice options, related to the $100 example above. As shown in Figure 10.4a, PT replaces the utility function u of *EU* theory with value function v, which is defined not over absolute outcomes (and resulting wealth levels) but in terms of relative gains or losses, i.e., as changes from a reference point, often the *status quo*. PT's value function maintains *EU*'s assumption that outcomes have decreasing effects as more is gained or lost (a property referred to by economists as "decreasing marginal sensitivity"). A person's degree of marginal sensitivity is measured by the parameter α in PT's power value function $v(x) = x^{\alpha}$. However, because outcomes are defined relative to a neutral reference point, the leveling off of increases in value as gains increase ("good things satiate") leads to a "concave" shape of the value function only in the domain of gains, as shown in Figure 10.4a. This concave shape is associated with risk-averse behavior, e.g., preferring the sure receipt of an amount much smaller than the expected value of a particular lottery over the opportunity to play the lottery. In contrast, the leveling off of increases in disutility as losses increase ("bad things lead to psychic numbing") leads to a "convex" shape of the value function in the domain of losses. This convex shape is associated with risk-seeking behavior, e.g., preferring a lottery of possible losses over the sure loss of an amount of money that is much smaller than the expected value of the lottery.

Another noteworthy characteristic of PT's value function is the asymmetry in the steepness of the function that evaluates losses and gains, with a much steeper function for losses ("losses loom larger"), also shown in Figure 10.4a. The ratio of the slope of the loss function over the slope of the gain function is referred to as *loss aversion*, and is another individual difference parameter, which is reflected by parameter λ. Empirical studies have consistently confirmed loss aversion as an important aspect of human choice behavior (Rabin, 1998; Camerer, 2005). It is also a likely explanation for real-world phenomena such as the endowment effect (Thaler, 1980), the *status quo* bias (Samuelson and Zeckhauser, 1988; Johnson and Goldstein, 2003), and the equity premium puzzle (Benartzi and Thaler, 1995), which describe behavior that deviates from the normative predictions of classical EU theory and risk–return models.

Just as PT suggests a subjective transformation of objective outcomes, it also suggests a psychological transformation of objective probabilities, p, into subjective decision weights, $\pi(p)$, which indicates the impact the event has on the decision. The original PT decision weight function, shown in Figure 10.4b, formalized empirical observations showing that small probability events receive more weight than they should, based on their likelihood of occurrence, while large probabilities receive too little weight. More recently, a more complex continuous function has been substituted (Tversky and Kahneman, 1992). See Chapter 11 of this volume for more details on PT.

PT also suggests that decision makers will simplify complex choices by ignoring small differences or eliminating common components of a choice. These editing process are not well understood, or easily captured by a formal model. While PT is significantly more complex than *EU*, its psychological modifications explain many anomalous observations that have accrued over many years. Risk–return models have also undergone similar psychological modifications in recent years (Sarin and Weber, 1993), and are discussed in 'Modeling decision making under uncertainty,' below.

DECISIONS UNDER UNCERTAINTY

The models of risk preference introduced above, will be more formally revisited in the following section. In this section, we examine some distinctions between different types of uncertainty and different ways of reducing or resolving uncertainty. In the process of doing so, we discuss recent suggestions that dual-processing systems are involved in risky choice.

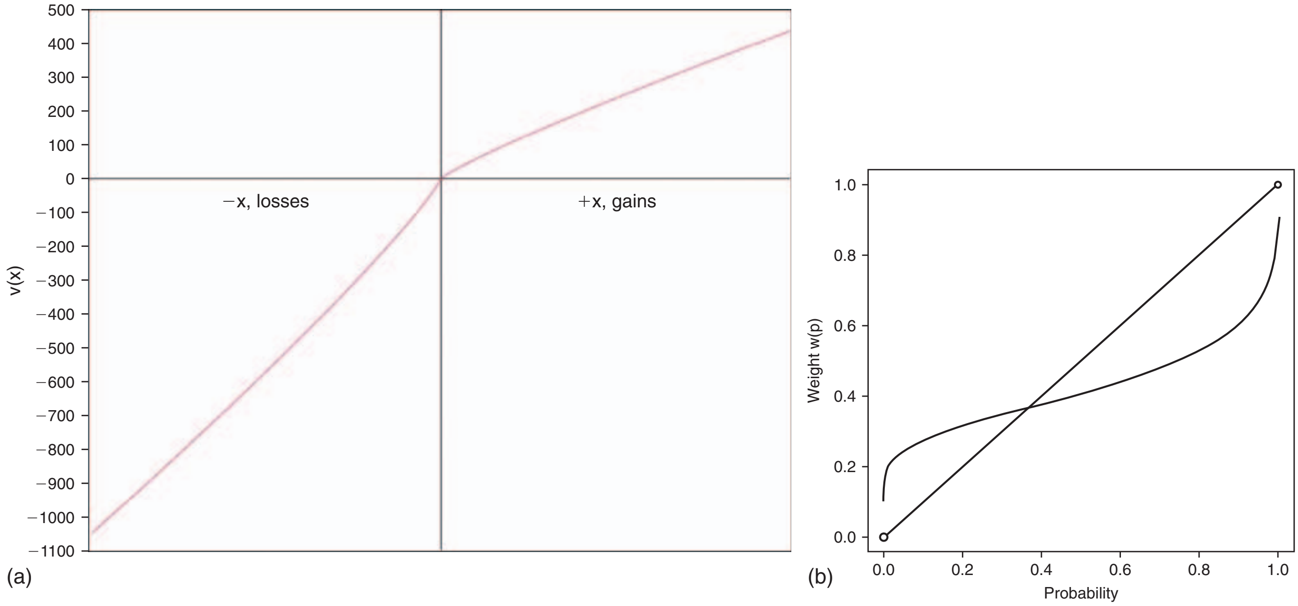


FIGURE 10.4 Prospect theory's (1979) value function (a) $v(x)$ which is $x^{0.88}$ for gains and $2.25\bullet x^{0.88}$ for losses, and (b) decision weight function $\pi(p)$.

Uncertainty

Types of Uncertainty

Benjamin Franklin famously stated that the only things certain in life are death and taxes. If anything, the amount of uncertainty in our world has increased between the eighteenth and twenty-first centuries. A common distinction is made between *aleatory* uncertainty, i.e., objective and irreducible uncertainty about future occurrences that is due to inherent stochasticity in physical or biological systems, and *epistemic* uncertainty, which is subjective and reducible, because it results from a lack of knowledge about the quantities or processes identified within a system. The uncertainty associated with the outcome of the toss of a coin is an everyday example of aleatory uncertainty, whereas not knowing the chlorine level of your swimming pool is an example of epistemic uncertainty. While epistemic uncertainty is reducible in principle, many domains may have limits to the precision of predicting events far into the future, due to the complex or chaotic nature of the processes that are giving rise to them (Lempert *et al.*, 2004). The social world provides uncertainties beyond those of the physical world, and game theory is a way of coping with the uncertainties that arise out of our limited ability to predict the behavior of others, as described in Chapters 5, 6, and 13 of this volume.

Degrees of Uncertainty

The economist Frank Knight was the first to make a conceptual distinction between decisions under *risk* and under *uncertainty* (1921: Ch. 7). *Risk* refers to situations where the decision maker knows with certainty the mathematical probabilities of possible outcomes of choice alternatives. *Uncertainty* refers to situations where the likelihood of different outcomes cannot be expressed with any mathematical precision. Rational-economic analysis assumes that uncertain situations can be reduced to risky situations. In the absence of any information about probabilities, all possible values (in the extreme, between 0 and 1) should be assumed to be equally likely, with the midpoint of the range of possible likelihoods (e.g., .5) as the best estimate, a line of reasoning referred to as the "ignorance prior." Contrary to this assumption, Ellsberg (1961) showed that people clearly distinguish between risky and uncertain options and have a clear preference for the former – a behavior that Ellsberg called *ambiguity aversion*[1].

Knowledge about the probability distribution of possible outcomes of a choice can lie anywhere on a continuum, from complete ignorance (not even the possible outcomes are known) at one end, through

---

[1] Some psychologists have argued that the word ambiguity ought to be reserved for situations that have a small number of possible interpretations – for example, the word "portfolio" referring to either a set of stocks held or to a set of artworks produced by a person. Situations that allow for a broad range of possible likelihoods of different events should be described as *vague*, and people's dislike of such situations as *vagueness aversion* (Budescu *et al.*, 1988; Budescu and Wallsten, 1995), though this change in terminology does not appear to have been adopted.

various degrees of partial ignorance (where outcomes may be known, but their probabilities not precisely specified, denoted as uncertainty or ambiguity), to risk (where the full outcome distribution is precisely specified), to certainty (where only a single, deterministic outcome is known to result).

Ambiguity aversion has been observed in both laboratory experiments and in real-world health, environmental, and negotiation contexts (see Curley and Yates, 1989; Hogarth and Kunreuther, 1989). While ambiguity aversion is a very stable phenomenon, it is not universally observed (Camerer and Weber, 1992). If the ambiguous choice option is in a domain in which the decision maker believes herself to have expertise, ambiguous options (e.g., sports bets) are often preferred to equivalent risky monetary lotteries (Fox and Tversky, 1995).

Ways of Resolving and Quantifying Uncertainty

Epistemic uncertainty can be resolved in different ways. People and other organisms learn in a number of different ways, as described in Part 4 of this volume. Personal experience powerfully affects memory and subsequent behavior: a single painful touch of a hot stove can prevent similar mishaps for a lifetime. Observational learning is an evolutionary innovation available only to humans, primates, and a few other species (Zentall *et al.*, 1988). Cultural learning, the ability to understand other's cautionary tales and anecdotes, extends the range of vicarious experience even further. Individuals who live in cooperative groups with the ability to communicate information in symbolic form can use the experience of others not just by direct observation, but also receive it in condensed form. The possible outcomes of investing in a particular company stock, for example, can be provided as a probability distribution of possible outcomes or as a time-series of past outcomes.

Multiple Processing Systems and the Resolution of Uncertainty

Clinical (Epstein, 1994), social (Chaiken and Trope, 1999), as well as cognitive psychologists (Sloman, 1996; Kahneman, 2003) have recently proposed very similar dual-processing models of decision making. Stanowich and West (1998) refer to the two hypothesized functional systems as "System 1" and System 2," others as rule-based or analytic versus experiential or associative systems. Which system is assumed to direct information processing in a given situation is often related to the way in which information about outcomes and their probabilities was acquired, over time from personal experience, or by external description (Erev and Barron, 2005).

Experiential processes correspond to the "concrete operations" described by Piaget (1962), while analytic processes are an example of his "formal operations," i.e., operations on ensembles of concrete experiences. Personal experience frequently contains strong feelings, making it memorable and therefore often dominant in processing (Loewenstein *et al.*, 2001; Slovic *et al.*, 2002). Strong feelings such as pleasure, pain, fear, and anger involve activation of a socio-emotional network of brain regions, in particular limbic and paralimbic structures, many of which are evolutionarily older than neocortical regions and found in all vertebrates (Cohen, 2005; Steinberg, 2007). By contrast, analytic processes that allow for planning, cognitive control, and self regulation involve prefrontal and parietal regions of the neocortex that have grown in size most in humans relative to other species (Cohen, 2005). The extent to which analytic processes occur in non-human animals is a subject of active investigation, though it seems clear that some processes, including those that underlie the syntactic structures of human language and the use of extended chains of logic, are uniquely human (Pinker, 1994).

Despite the current popularity of these dual-process explanations, not too strong a separation should be drawn between experiential and analytic processing (Keysers *et al.*, 2008). Even simple reflexes can be influenced by neocortical processes, and analytic reasoning can lead to strong feelings. A given decision always involves and integrates both kinds of processes. The role of analytic processes in the understanding of uncertainty and in decisions involving such information has, however, often been overestimated, and the role of experiential processes has until recently not been sufficiently appreciated (Loewenstein *et al.*, 2001).

Earlier in the chapter we discussed different ways in which human decision makers can resolve epistemic uncertainty, from personal trial-and-error learning from the feedback provided by repeated sampling of available choice alternatives to the (external) provision of a numeric or graphic probability distribution of possible outcomes. The first of these ways has recently been labeled *decisions from experience*, and the second *decisions from description* (Hertwig *et al.*, 2004; Weber *et al.*, 2004). Research on decisions under these two ways of becoming knowledgeable about outcome distributions has historically been conducted in parallel by different research communities, with empirical research on human decision making virtually exclusively employing decisions from description, and empirical research on animal learning and decision making under uncertainty by necessity employing decisions from experience. Direct comparisons of choices under these two learning conditions in humans suggest that choices differ when small probability events are involved[2]. While rare events get more weight than they deserve by their probability of occurrence in decisions from description as modeled by PT's probability weighting function (Tversky and Kahneman, 1992), they tend to be underweighted in decisions from experience, unless they have recently occurred, in which case they are hugely overweighted (Weber, 2006). For more information on model differences and empirical results, see Weber *et al*. (2004).

MODELING DECISION MAKING UNDER UNCERTAINTY

In this section we revisit the two models introduced in their historical context at the beginning of this chapter, with the goal of showing how descriptive models of risky choice have built on them. Since *EU* theory and PT are described elsewhere (see Chapters 3 and 11 of this volume), we focus only on their general features and their commonalities to prescriptive and descriptive risk–return models of risky choice.

Risk-Taking and Risk Attitudes in EU and PT

Not All Apparent Risk-taking May be Due to Risk Attitude

Both the EU and the traditional risk–return approach to risky decision making model differences in choice behavior with a single parameter, referred to as "risk attitude" or "risk tolerance." This parameter simply describes the curvature of the utility function or the slope of the risk–return tradeoff, and is identified empirically from a person's choices. For example, someone who is indifferent between \$45 for sure and a 50/50 gamble between \$0 and \$100 is risk averse. The \$5 difference between the EV of the gamble (i.e., \$50) and the certainty equivalent of \$45 is referred to as the risk premium. Greater risk aversion results in a larger risk premium. The label "risk attitude" suggests that such behavior is motivated by an attitude, typically a stable construct, i.e., a personality trait.

Unfortunately for the interpretation of risk attitude as a personality trait, risk-taking is far from stable across situations for most individuals (Bromiley and Curley, 1992). The same person often shows different degrees of risk-taking in financial, career, health and safety, ethical, recreational, and social decisions (MacCrimmon and Wehrung, 1986; Weber *et al*., 2002; Hanoch *et al*., 2006). This leaves two options. Either there is no stable individual difference in people's attitude towards risk, contrary to the intuition that people differ on this dimension, or we need to find a way to measure risk attitude in a way that shows stability across domains by factoring out other (more situationally determined) contributors to apparent risk-taking.

Constant and Relative Risk Aversion in EU

EU explains the fact that people's certainty equivalents for lotteries typically are below the lotteries' EV by a concave function that turns objective amounts of money into their utility equivalent, with increasing amounts of money generating increased utility (positive slope, i.e., a positive first derivative), but less and less so (i.e., a negative second derivative). There is a large number of functions that have this general characteristic, not just the power function shown in Figure 10.2. Economists Kenneth Arrow and James Pratt thus tried to derive some measures of risk aversion independent of the utility function's functional form. They did so by linking risk aversion and the risk premium described above and, in particular, defined two indices that specified how a person's risk-taking would change as her wealth increases. There being more detail in Chapter 3 of this volume, we will only describe two types of effects here. The Arrow–Pratt (1964) measure of absolute risk aversion, defined as:

$$\mathrm{ARA}_u(x) = -u''(x)/u'(x) \tag{10.4}$$

where u' and u'' denote the first and second derivative of utility function u, specifies the absolute value of the risk premium associated with a given lottery. As shown in Figure 10.5 (left column), exponential utility functions have the property of constant absolute risk aversion (CARA), meaning that the decision maker would pay the same risk premium to avoid the uncertainty of a given lottery (e.g., \$5 for the 50/50 lottery between \$100 or nothing) at all levels of wealth. Arrow (1965) more realistically assumed that most people show decreasing absolute risk aversion, i.e., would be more likely to play the gamble at higher levels of wealth, and thus pay a smaller risk premium to avoid it.

---

[2] Differences in prediction between choices made by people under description (based on PT) vs under experience (based on reinforcement learning models like the Fractional Adjustment Model) start to occur when risky options contain probabilities less than .25, and tend to get larger the smaller the probabilities of some outcomes.

The other Arrow–Pratt measure, relative risk aversion, defined as:

$$\mathrm{RRA}_u(x) = -(x\,u''(x))/u'(x) \qquad (10.5)$$

specifies the percentage value of wealth the *EU* maximizer is willing to put at risk. As shown in Figure 10.5 (right column), power utility functions have the property of constant relative risk aversion (CRRA), meaning that the decision maker is willing to put the same percentage of wealth at risk (e.g., 40% in Figure 10.5) at all levels of wealth. Arrow (1965) assumed that instead, most people would show increasing relative risk aversion.

Accounting for Domain Differences in Risk-taking

An early attempt to restore cross-situational consistency to the construct of risk attitude argued that utility functions derived from risky choices, $u(x)$, consist of two components, one measuring the (typically decreasing) marginal value ($v(x)$) of the outcome dimension (e.g., two bananas not being twice as rewarding as one banana), the other measuring the (typically averse) attitude towards risk, $u(v(x))$, which reflects a dislike of the fact that in a lottery one does not know for sure what one will get, resulting in the risk premium discussed above. In such cases, $u(v(x))$ is not as large as $v(x)$, and gets increasingly smaller the more $v(x)$ is at stake. If the index of the curvature of risky utility functions is the sum of these two contributions, then domain differences in curvature could be the result of the different marginal values for different outcomes dimension (e.g., the incremental value of an additional dollar vs the incremental value of an additional life saved), while the true attitude towards the risk or uncertainty with which these outcomes were obtained could be the same across domains. Figure 10.6 provides an example from a hypothetical person who has decreasing marginal value for additional bananas (shown in the top left panel) and slightly increasing marginal value for additional glasses of wine. As indicated in the middle panels, by the straight line that maps marginal value into utility, this person happens to have a completely neutral attitude towards risk, i.e., her anticipated enjoyment of bananas or glasses of wine is the same, regardless of whether these are acquired for certain or as part of

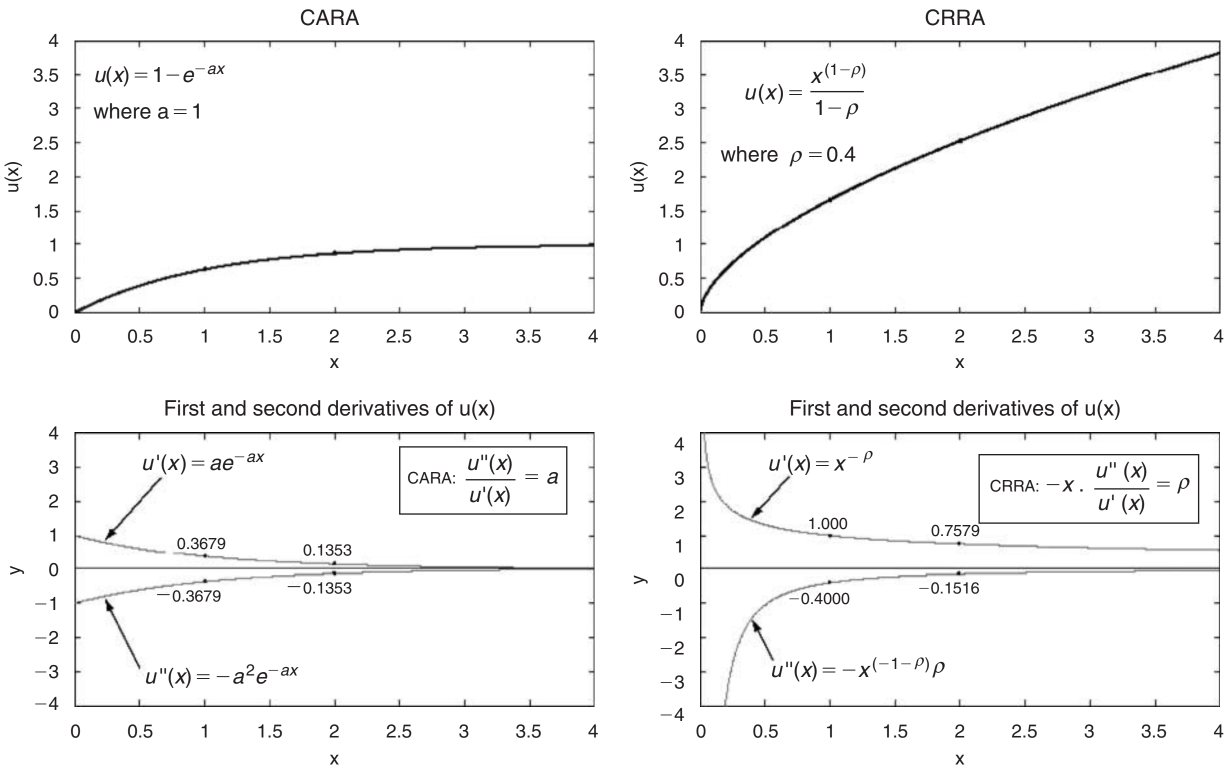


FIGURE 10.5 Constant absolute risk aversion (CARA, left column) and constant relative risk aversion (CRRA, right column). The top panel shows the described utility function, the bottom panel its first and second derivative.

a lottery. Because of the difference in marginal value, however, a utility function inferred from risky choices will show her to be risk averse for bananas (bottom left panel) but risk-seeking for glasses of wine (bottom right panel). Dyer and Sarin (1982) suggested that possible domain differences in riskless marginal value be factored out of an assessment of risk attitude, and thus replaced the Arrow–Pratt (1964) measure of ARA with what they referred to as *relative risk attitude*:

$$-u''(v(x))/u'(v(x)) \tag{10.6}$$

where $v(x)$ denotes the riskless marginal value function. When Keller (1985) compared people's

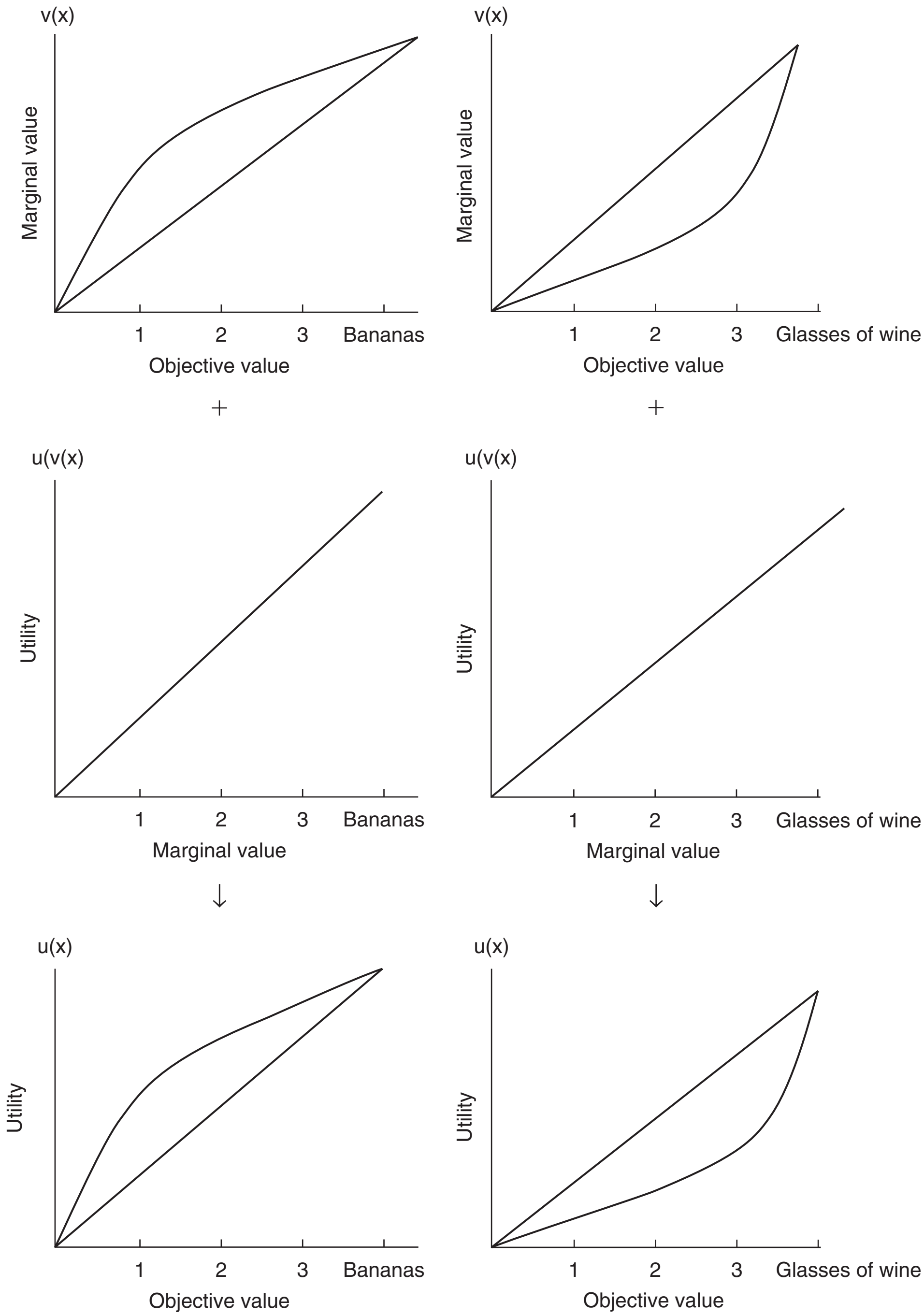


FIGURE 10.6 Decomposition of utility function $u(x)$ (bottom row) into marginal value function $v(x)$ (top row) and attitude towards risk function $u(v(x))$ (middle row).

Arrow–Pratt measure of risk attitude (inferred from risky choices in various decision domains) to their relative risk attitudes (inferred from choices and marginal value functions in the same domains), she found that the two agreed in only a small number of cases, supporting the usefulness of unconfounding attitude towards uncertainty from non-linear marginal value. Unfortunately, relative risk attitudes did not show any more consistency across decision domains for any given respondent than the Arrow–Pratt ARA measure.

PT does not directly address the issue of inconsistency in risk-taking in different decision domains, but suggests other reasons we might see different risk-taking. Because a reference point divides outcomes into relative gains and relative losses, decreasing marginal utility produces a concave function and thus risk-averse choice for gains, but a convex function and thus risk-seeking choices for losses. In addition, the loss function has a steeper slope than the gain function (loss aversion), and probability weighting is non-linear. Thus PT, to the extent it is a descriptive theory of choice, suggests many reasons why risk-taking many seem unstable: First, the representation of the problem might change reference points, changing the apparent risk attitude. Second, to the extent that a person's degree of loss aversion differs for outcomes in different domains, prospect theory could account for domain differences in risk-taking. Gaechter *et al*. (2007) provide evidence that loss aversion can differ for different attributes, in their case as a function of attribute importance and the decision maker's expertise in the domain.

Behavioral extensions of risk-return models (Sarin and Weber, 1993) account for domain differences in risk-taking by questioning the equating of return with *EV* and of risk with outcome variance. While studies of financial decisions typically find that the *EV* of risky investment options presented in decisions from description is a good approximation of expected returns (Weber *et al*., 2005), survey data assessed in populations known to differ in actual risk-taking behavior suggest that risk-takers judge the expected benefits of risky choice options to be higher than do control groups (Hanoch *et al*., 2006). A large and growing literature has also examined perceptions of risk, both directly (by assessing people's judgments or rankings of the riskiness of risky options and modeling these, often on an axiomatic basis) and indirectly (trying to infer the best-fitting metric of riskiness from observed choices under the assumption of risk–return tradeoffs) (see Weber, 2001a, for further details). These studies are unanimous in their verdict that the variance or standard deviation of outcomes fails to account for perceived risk, for a variety of reasons. First, deviations above and below the mean contribute symmetrically to the mathematically defined variance, whereas perceptions of riskiness tend to be affected far more by downside variation (e.g., Luce and Weber, 1986). Second, variability in outcomes is perceived relative to average returns – a standard deviation of ±$100 is huge for a risky option with a mean return of $50, but amounts to rounding error for a risky option with a mean return of $1 million. The coefficient of variation (CV), defined as the standard deviation (SD) that has been standardized by dividing by the *EV*:

$$\mathrm{CV}(X) = \mathrm{SD}(X)/\mathrm{EV}(X), \tag{10.7}$$

provides a relative measure of risk, i.e., risk per unit of return. It is used in many applied domains, and provides a vastly superior fit to the risk-taking data of foraging animals and people who make decisions from experience (Weber *et al*., 2004). Weber *et al*. (2004) show that simple reinforcement learning models that describe choices in such learning environments predict behavior that is proportional to the CV and not the variance. Kacelnik and colleagues have explained animal risk-taking that is proportional to the CV, using a model called Scalar Utility Theory, which postulates that the cognitive representation of outcomes follows Weber's Law (1834) – namely, that the spread of the distribution of expected outcomes is proportional to its mean (see, for example, Marsh and Kacelnik, 2002).

Finally, affective (i.e., non-rational or non-consequential) responses to risky situations have been shown to play a large role in both the perception of the riskiness of risky choice options and in risky choice. The greater volatility in responses observed in decisions from experience relative to decisions from description, for example, where behavior is influenced more by more recent experiences[3] can be seen as resulting from the salience of emotional reactions to recent outcomes. Familiarity with risky choice options or a risky choice domain lowers the perceptions of the choice options' riskiness[4]. The home bias effect in investing, i.e., the tendency to invest a larger than prudent amount of one's assets into stocks in one's home country or into stock of the company one works for, has been shown to be mediated by perceptions of lower risk of familiar investment opportunities (Weber, 2006).

How to Measure Risk Attitude

The behavioral research we reviewed strongly suggests that there is no single measure of "risk attitude"

---

[3] An adaptive learning rule in non-stationary environments.

[4] In evolutionary times, safer options provided longer periods of survival, with longer opportunities to acquire familiarity with choice options.

that can be inferred from observed levels of risk-taking. To find a person's true attitude towards risk (liking it for its excitement vs disliking it for the anxiety it induces) requires that we decompose observed risk-taking into the multiple factors (including risk attitude) that influence it. We would like to suggest that EU-based and other measures that simply re-describe the observed level of risk-taking (where greater risk-taking is typically operationalized as choosing options that have greater variance, while controlling for EV) use the term "risk-taking" instead.

One criterion for deciding how to assess individual differences in risky choice behavior is the purpose of the assessment, which usually falls into one of the following two categories: prediction or intervention. When measuring levels of risk-taking with the objective of predicting risk-taking in other situations, it is important to use a decision task that is as similar as possible to the situation for which behavior is being predicted. Given what we know about the domain-specificity and sign-dependence of risk-taking, assessment questions should come from the same domain and match the target situation in other respects. Weber *et al.* (2002) found that assessed risk-taking for monetary gambling decisions predicted real-world investment decisions far worse than assessed risk-taking for investment decisions, even though both were about monetary returns. Nosic and Weber (2007) confirmed that risk-taking for stock investments was not related to risk-taking for money lotteries, but was predicted by risk attitude, risk perception, and perceptions about return elicited in a stock-related context. It is thus not surprising that risk-taking indices like the level of relative risk aversion measure inferred by Holt and Laury (2001) from gambling choices, while widely used, have had only very mixed results in predicting risk-taking in other domains.

When intervention is the goal of efforts to assess individual differences in risk-taking (e.g., to make women less risk averse in their financial investment decisions), it becomes important to understand the causes of the apparent risk aversion or risk-seeking at a process level. One needs to understand whether apparently risk-averse decisions are driven by gender-specific differences in true attitude towards risk (e.g., women assessing risks and returns accurately, but disliking the risks more than men do), or whether other differences lie at the root of the gender differences in behavior (for example, differences in the subjective perception of risks or benefits, or differences in loss aversion). A more fine-grained assessment of determinants of risk-taking becomes important, because different causes of the behavior will dictate different interventions if seeking to effect change.

Risk-Taking and Risk Attitude in Psychological Risk–Return Models

Psychophysics, the study of the relationship between physical stimuli and their subjective perception, was the first topic of investigation of scientific psychology. The observed mappings between physical outcome dimensions (decibels) and subjective perception (loudness) were found to be not only non-linear, but also subject to context effects (see Weber, 2004). With the argument that similar non-linear and complex transformations might map objective outcome variation into perceived risk, and objective outcome EV into expected benefits, researchers from several disciplines (see Sarin and Weber, 1992) have recently generalized the normative finance risk–return model to allow for subjective perception of risks and returns which are, as before, traded off to determine willingness to pay (WTP) for risky option X:

$$\mathrm{WTP}(X) = V(X) - bR(X) \tag{10.8}$$

In these generalized psychophysical risk–return models, all three components, $V(X)$, $R(X)$, and tradeoff parameter b, are psychological variables, which can differ as the result of individual or situational characteristics. Behavioral evidence shows that the same objective outcome variation can be perceived in systematically different ways by different individuals and cultures (Brachinger and Weber, 1997; Weber, 2001a, 2001b). The characteristic that differentiates entrepreneurs from other managers, for example, contrary to managerial folklore, is *not* a more positive *attitude* towards risk, but instead an overly optimistic *perception* of the risks involved (Cooper *et al.*, 1988). For outside observers who perceive risk more realistically, entrepreneurs will appear to take great risk; however, when differences in risk perception are factored out, entrepreneurs – just like other managers – demonstrate a preference for tasks that they see as only moderate in risk (Brockhaus, 1982).

When perceived risk and return replace the statistical moments of variance and EV in the prediction equation of risk-taking, the tradeoff coefficient b can be interpreted as an index of true attitude towards risk. Labeled *perceived risk attitude* (PRA) by Weber and Milliman (1997), it is a measure of the degree to which individuals find perceived risk attractive (or unattractive) and therefore will choose alternatives that carry greater (or less) risk, all other things being equal. Weber and Hsee (1998) obtained risk judgments as well as minimum buying prices for risky financial investment options from decision makers in the USA, Germany, the People's Republic of China, and Poland.

Both risk judgments and buying prices showed significant cross-national differences, with Americans perceiving the highest risks and Chinese paying the highest prices. However, after differences in risk perception were taken into consideration, the proportion of individuals who were perceived-risk averse or perceived-risk seeking were not significantly different in the four countries, with the majority being perceived-risk averse, and only a small percentage in each country being perceived-risk seeking.

Some psychologists have questioned the assumption of finance models that people will and should strive to minimize risk, arguing instead that people's ideal point for risk or uncertainty could differ, either as a personality difference (Lopes, 1987) or as a situational difference (Weber and Kirsner, 1997). Ideal-point models (Coombs, 1975) assume a person will perceive the riskiness of an alternative as the deviation between the alternative's level of uncertainty and the person's ideal point on the uncertainty continuum. Perceived risk of an alternative with a high objective level of uncertainty would be high for a person with a low ideal point, but low(er) for a person with a high ideal point. Individual differences in ideal points for risk and uncertainty have been measured by the construct of sensation-seeking (Zuckerman, 1979) which has a biological basis (Zuckerman *et al*., 1988) and varies with age and gender. Bromiley and Curley (1992) report evidence linking sensation-seeking to behavioral correlates that include greater risk-taking, especially in the health/safety and recreational domain. Weber *et al*. (2002) also report high positive correlations between sensation-seeking and its subscales in several content domains, with especially high correlations between the thrill-and-adventure-seeking subscale and recreational risk-taking, and the disinhibition subscale and ethical risk-taking. Consistent with the predictions of ideal-point models, the path by which differences in sensation-seeking seem to affect risk-taking appears to be differences in the perceptions of risk and of benefits, rather than differences in attitude towards perceived-risk. In other words, groups known for high levels of sensation-seeking (e.g., bungee jumpers or teenage boys) seem to take large risks because they perceive the risk to be smaller or the benefits to be larger than do other groups, and not because they cherish (perceived) risk to a greater extent (Hanoch *et al*., 2006).

Process-Tracing Methods and Process Data

Cognitive psychology has long tested models of risky choice using process analysis. The basic idea is to test models not just by their outputs, but also by their use of inputs and intermediate products. For example, EU models suggest that outcomes are weighted by their probability of occurrence. For process analysis, this suggests that decision makers would look at each payoff and its probability *within* a choice alternative in close temporal order. In contrast, models that emphasize anticipated regret suggest that comparisons of the outcomes of different choice options for the same states of the world are relevant to choice, thus making a different prediction for information search from EU, namely a significant number of comparisons of pairs of outcomes *between* alternatives.

A wide variety of process-analysis techniques exist, including asking people to talk aloud as they make risky choices (see Ericsson and Simon (1993) for a review of the method, and Bettman and Park (1980) for an example). Information acquisition while making a choice has been examined by recording eye fixations on visually displayed information (Russo and Dosher, 1983) or, when using a computer to make a decision, by recording the mouse clicks that reveal information on the computer screen (Payne *et al*., 1993; Costa-Gomes *et al*., 2001; Gabaix *et al*., 2006; Johnson *et al*., 2008). In many ways, process analysis is a close relative of brain-imaging data, since the goal is to add other sources of data that inform models, and to provide additional constraints on theories.

When these techniques have been applied to risky choice, several facts emerge. First, complex displays (for example, lotteries with many outcomes or choices between many risky options) produce a different kind of processing than do simple choices between two binary lotteries. As posited by the editing phase of PT, when faced with complex displays or time pressure, decision makers try to eliminate options and attend to only a subset of the available information. This suggests that imaging studies of risky choice that typically use very simple stimuli will speak to different processes than those used in more complex environments.

Second, even with simple choices, different ways of measuring preferences can invoke different choice processes. Consider the classic behavioral observations of inconsistent preferences across response modes, *preference reversals*. Here, people will choose one gamble over another but then, when asked to price the same two gambles, will give a lower price to the one they chose. These reversals (Lichtenstein and Slovic, 1971) provide a major source of evidence that EU is an incomplete model of decision making under risk. Process data suggest that these reversals occur because people use different processes and put different weight on probabilities and payoffs when generating a price than when making a choice (Schkade and Johnson, 1989; Mellers *et al.*, 1992). Observed preferences are not (just) an expression of inherent preferences; they also depend on the processes used to generate and express the preference.

Third, studies of choices between pairs of simple gambles tend to show some support for accounts that posit the weighting of outcomes by probabilities, consistent with EU and PT. While Brandstaetter *et al.* (2006) argue that a heuristic model, called the priority heuristic, that makes different and simpler comparisons than PT accounts for the same observed choices as PT, process-tracing studies show substantial inconsistencies with their heuristic model at a process level (Johnson *et al.*, 2008).

Finally, there are marked individual differences in processes used to make risky choices. No single process seems to be used by all people, and there is significant evidence of shifts in strategies across different kinds of problems (Ford *et al.*, 1989). In addition, there are strategy shifts when factors such as the time available to make a decision or the nature of the choice set changes (Ben Zur and Breznitz, 1981; Payne *et al.*, 1988).

Neuroimaging Studies and Data

Neuroimaging techniques have added to our understanding of risky decision making by providing evidence that hypothesized psychological processes and individual and situational differences in such processes have physical manifestations in brain processes. While this may seem obvious and unremarkable to some, it allows us to settle some long-standing arguments between psychologists and economists about the equivalence of different stimulus presentations, decision situations, or prior learning conditions.

While the correct interpretation of both behavioral and neural results is not uncontroversial, comparisons of brain activation of people who choose between choice options that involve ambiguous outcomes vs choice options that involve the equivalent risky outcomes suggest that these two choice situations differ, and how (Hsu *et al.*, 2005; Huettel *et al.*, 2006). Neuroimaging studies suggest that there is strong path dependence in the brain's reaction to economic quantities like likelihood or risk/variance. While normative economic models do not distinguish between knowledge about the likelihood of different consequences that was acquired either by trial-and-error learning or by being given a statistical summary, as long as the accuracy of knowledge and source credibility are controlled for, psychological models make different predictions for decisions from experience and decisions from description, and both process-tracing methodologies and neuroimaging data can be used to validate these psychological accounts (Delgado *et al.*, 2005). While it does not matter, to finance models of risk-taking, whether the expected value and variance of risky choice options is manipulated in a given choice set by varying the probabilities of different outcomes or their magnitudes (or both), neuroimaging studies that look at the effect of EV and variance on risk-taking tend to observe very different patterns of activation based on such differences in manipulation (Preuschoff *et al.*, 2006 vs Figner *et al.*, 2007; also see Chapter 23 of this volume). Studies that have examined brain activation in response to gains vs losses, looking for the neural equivalent of loss aversion, also find different patterns of brain activation depending on whether each decision is resolved or not (Tom *et al.*, 2007 vs Huettel *et al.*, 2006), or whether people make decisions or just contemplate the options (Breiter *et al.*, 2001).

SUMMARY AND IMPLICATIONS

Psychological and neuroscience studies of risk-taking have identified a wide range of factors, some exogenous and some endogenous, that influence risk-taking, as reviewed in this chapter. Multiple processes (some more effortful and analytic, others automatic, associative, and often emotion-based) are in play when a preference between different risky options is constructed. As decision makers with limited attention and processing capacity, we need to be selective in what information we use, and have to find shortcuts to process it. Situational characteristics, like the way in which information about choice options is presented to us, or the nature of the task (e.g., choice vs a price judgment), influence risk-taking by focusing our attention on different subsets of information (e.g., the magnitude of outcomes for price judgments, their probabilities for choices) or by facilitating different relative comparisons in our search for the better option. Characteristics of the decision maker (e.g., gender) often interact with characteristics of the situation (e.g., the domain of the decision) in determining risk-taking. This is either because different decision makers use different processes to different degrees (e.g., decision makers with greater cognitive capacity can make more use of effortful analytic processes – see Chapter 4 of this volume) or because the same processes result in different output (e.g., decision makers familiar with a choice domain may experience positive emotions such as comfort or confidence when contemplating risky options in that domain, whereas decision makers unfamiliar with the domain will experience negative emotions such as anxiety (Weber *et al.*, 2005).

Figure 10.7 summarizes the implications of this chapter's review of the multiple determinants of risk preference for the frequently asked question: How can or should I assess the risk attitudes of a given group of decision makers? As the flowchart indicates, the first diagnostic question that needs to be answered in such situations is: Why are we assessing risk; what is the purpose of the desired assessment? If the purpose is simply to predict what decision makers will do in (another) risky choice situation, the reasons for observed risk-taking need not be investigated. The main concern in such a predictive risk-taking

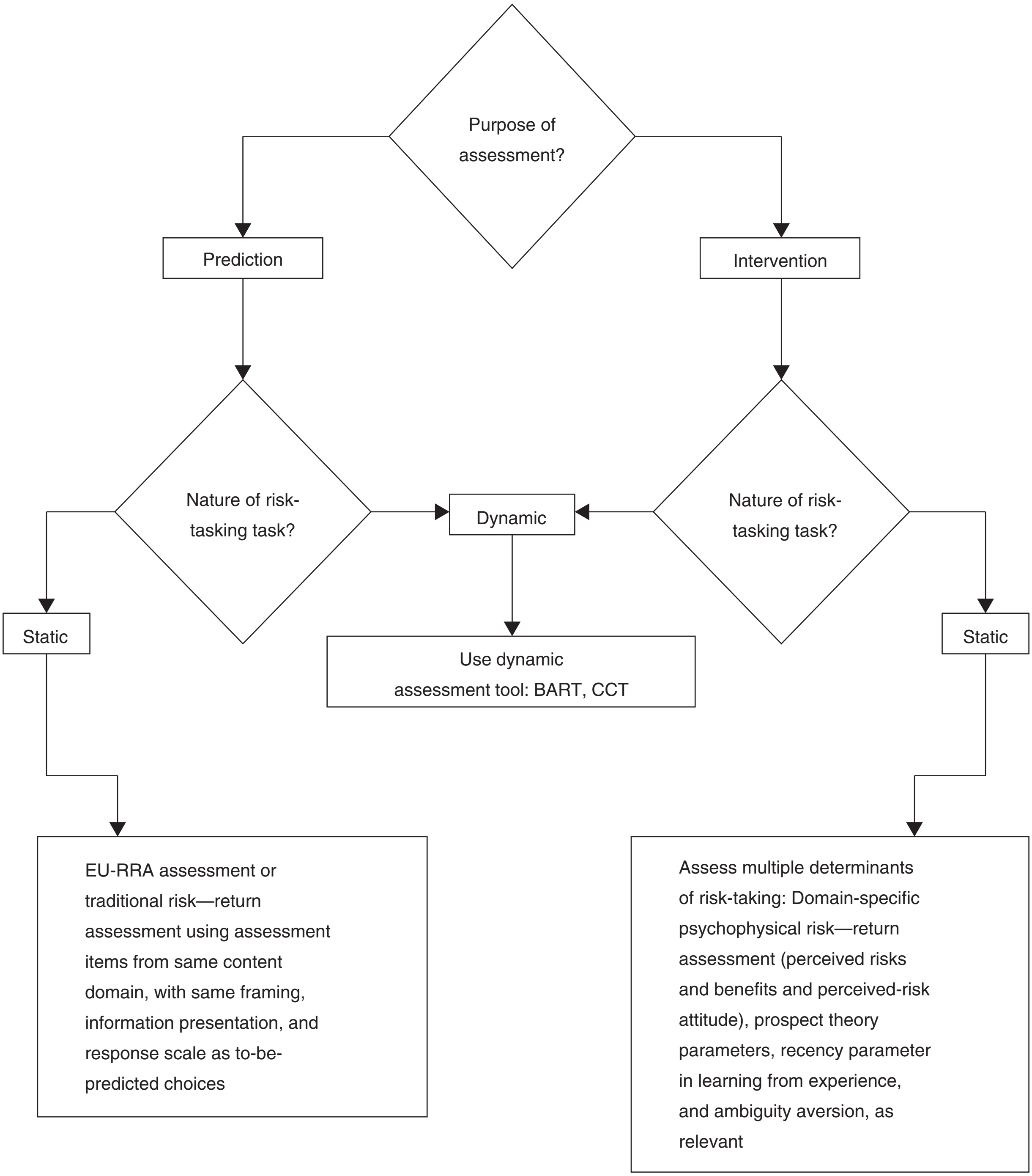


FIGURE 10.7 Decision tree for assessment of risk attitude.

assessment is to use a methodology that has high fidelity to the future risk-taking situation to which the predictive assessment will be applied. As discussed earlier in the chapter, risk-taking is often domain specific, which makes ostensibly "content-free" utility assessment tools like the Holt and Laury (2002) lotteries better predictors of risk-taking in monetary gambling choices than in risky agricultural production decisions. Weber *et al*. (2002) found that the gambling subscale of their Domain Specific Risk Taking (DOSPERT) scale was a significantly better predictor of self-reported gambling behavior than even of monetary investment decisions. This suggests that it is best to use domain-specific risk attitude assessment tools or to "translate" tools like the Holt and Laury lotteries into the domain context in which one is trying to predict.

Another important component of looking for a high-fidelity match between assessment tool and application is the nature of the risk-taking behavior that one is trying to predict. Much real-world risk-taking is incremental and dynamic, involving sequential risk-taking with feedback, from taking risks in traffic to risky substance (ab)use. Given what we have learned about the susceptibility of neural processing of risky decision situations to learning and feedback, it should come as no surprise that risk-taking in such dynamic contexts is typically not predicted by static assessment tasks, like one-shot lottery choices that are not resolved until the end of the assessment (Wallsten *et al*., 2005). If the risk-taking to be predicted is dynamic, dynamic task assessment tools like the Balloon Analogue Risk Task (BART; Lejuez *et al*., 2002) or the diagnostically more sophisticated Columbia Card Task (CCT; Figner *et al*., 2007) should be employed. These dynamic assessment tools come closer to repeated real-world investment or gambling decisions, in which previous outcomes often influence subsequent gambling or investment behavior, leading to such phenomena as gambling with house money (Thaler and Johnson, 1992), or escalation of commitment (Weber and Zuchel, 2005).

Even for static risk-taking applications, task and choice set differences often influence risk-taking behavior and thus should be controlled for by making the assessment tool similar to the target situation in those respects. Apparent risk-taking has been shown to vary when preferences between risky options are expressed in different ways, e.g., by choices vs bids vs buying prices vs selling prices (Lichtenstein and Slovic, 1971; Holt and Laury, 2002). Since gain vs loss framing of choice options and the way decision makers have learned about outcome distributions affect risky choice, these variables should also be equated between the assessment and the to-be-predicted task.

Recommendations for assessment procedures get even more complicated for the right path in Figure 10.7, when the goal of the assessment is some intervention to change risk-taking in a target group of decision makers. In these situations, we need to determine the cause(s) of taking more or less risk than is normatively desirable, because different causes call for different interventions. Researchers may have some hypothesis about the underlying cause (e.g., an inappropriate attitude towards risk), but this diagnosis needs to be established by assessments that (1) measure the construct "risk attitude" in ways that are not confounded with other possible causes, and (2) rule out competing diagnoses. Inferring an index of risk aversion based on some assumed functional form for utility from a set of choices simply will not suffice, as discussed previously. Rather than assessing a single parameter (absolute or relative risk aversion) from such choices, at the very least the three individual difference parameters of PT should be assessed, to determine whether loss aversion or distortions in probability weighting contribute to the observed behavior or whether it is only due a decreasing marginal utility or value. In addition, decision-makers' perceptions of a choice option's risks and returns can be assessed and evaluated for accuracy. Regressing observed preference (e.g., willingness to pay) for risky options on perceptions of risks and returns allows for an assessment of true risk attitude, i.e., positive or negative reaction to risk as it is perceived.

While the behavior of people in situations of risk and uncertainty is complex and multiply determined, the broader set of tools provided by a psychological and neuroeconomic understanding of risk preference allows for a far more nuanced assessment and understanding of both general behavior patterns and individual or group differences in behavior. To the extent that psychology and neuroscience help explain the departures from the normative models described by economics, better interventions can be developed. Given the importance of accurate predictions of risk preference and of effective interventions to modify socially undesirable levels of risk-taking, we expect that the success of neuroeconomic methods will significantly contribute to greater acceptance of behavioral models by traditional economics.

Acknowledgments

Preparation of this chapter was facilitated by fellowships at the Russell Sage Foundation for both authors, National Institute of Aging grant 1R01AG027934-01A2, and National Science Foundation grant SES-0720932.

References

Allais, P.M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école americaine. *Econometrica* 21, 503–546.

Arrow, K.J. (1965). *Aspects of the Theory of Risk-Bearing*. Helsinki: Yrjö Hahnsson Foundation.

Bell, D.E. (1982). Regret in decision making under uncertainty. *Operations Res.* 30, 961–981.

Benartzi, S. and Thaler, R.H. (1995). Myopic loss aversion and the equity premium puzzle. *Q. J. Economics* 110, 73–92.

Ben Zur, H. and Breznitz, S.J. (1981). The effects of time pressure on risky choice behavior. *Acta Psychologica* 47, 89–104.

Bernoulli, D. (1954/1738). Exposition of a new theory on the measurement of risk. , [translation by L. Sommer of D. Bernoulli, 1738, Specimen theoriae novae de mensura sortis, *Papers of the Imperial Academy of Science of Saint Peterburg* 5, 175–192. *Econometrica* 22, 23–36.

Bettman, J.R. and Park, C.W. (1980). Effects of prior knowledge and experience and phase of the choice process on consumer decision-processes: a protocol analysis. *J. Consumer Res.* 7, 234–248.

Bodie, Z. and Merton, R.C. (1999). *Finance*. Englewood Cliffs, NJ: Prentice Hall.

Brachinger, H.W. and Weber, M. (1997). Risk as a primitive: a survey of measures of perceived risk. *OR Spektrum* 19, 235–250.

Brandstätter, E., Gigerenzer, G., and Hertwig, R. (2006). The priority heuristic: making choices without trade-offs. *Psychological Rev.* 113, 409–432.

Braun, M. and Muermann, A. (2004). The impact of regret on the demand for insurance. *J. Risk Ins.* 71, 737–767.

Breiter, H.C., Aharon, I., Kahneman, D. *et al*. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 21, 619–639.

Brockhaus, R.H. (1982). The psychology of the entrepreneur. In: C.A. Kent, D.L. Sexton, and K.G. Vesper (eds), *The Encyclopedia of Entrepreneurship*. Englewood Cliffs, NJ: Prentice Hall.

Bromiley, P. and Curley, S.P. (1992). Individual differences in risk-taking. In: J.F. Yates (ed.), *Risk-taking Behavior*. New York, NY: Wiley, pp. 87–132.

Budescu, D.V. and Wallsten, T.S. (1995). Processing linguistic probabilities: general principles and empirical evidence. In: J.R. Busemeyer, R. Hastie, and D. Medin (eds), *The Psychology of Learning and Motivation: Decision Making from the Perspective of Cognitive Psychology*. New York, NY: Academic Press, pp. 275–318.

Budescu, D.V., Weinberg, S., and Wallsten, T.S. (1988). Decisions based on numerically and verbally expressed uncertainties. *J. Exp. Psychol. Hum. Percept. Perf.* 14, 281–294.

Camerer, C. (2000). Prospect theory in the wild. In: D. Kahneman and A. Tversky (eds), *Choice, Values, and Frames*. New York, NY: Cambridge University Press, pp. 288–300.

Camerer, C. (2005). Three cheers – psychological, theoretical, empirical – for loss aversion. *J. Marketing Res.* 42, 129–133.

Camerer, C. and Weber, M. (1992). Recent developments in modeling preferences: uncertainty and ambiguity. *J. Risk Uncertainty* 5, 325–370.

Chaiken, S. and Trope, Y. (1999). *Dual-process theories in social psychology*. New York, NY: Guilford Press.

Cohen, J.D. (2005). The vulcanization of the human brain: a neural perspective on interactions between cognition and emotion. *J. Econ. Persp.* 19, 3–24.

Coombs, C.H. (1975). Portfolio theory and the measurement of risk. In: M.F. Kaplan and S. Schwartz (eds), *Human Judgment and Decision*. New York, NY: Academic Press, pp. 63–68.

Cooper, A.C., Woo, C.Y., and Dunkelberg, W.C. (1988). Entrepreneurs' perceived chances for success. *J. Business Vent.* 3, 97–108.

Costa-Gomes, M., Crawford, V.P., and Broseta, B. (2001). Cognition and behavior in normal-form games: an experimental study. *Econometrica* 69, 1193–1235.

Curley, S.P. and Yates, J.F. (1989). An empirical evaluation of descriptive models of ambiguity reactions in choice situations. *J. Math. Psychol.* 33, 397–427.

Delgado, M.R., Miller, M.M., Inati, S., and Phelps, E.A. (2005). An fMRI study of reward-related probability learning. *NeuroImage* 24, 862–873.

Dyer, J. and Sarin, R. (1982). Relative risk aversion. *Management Sci.* 28, 875–886.

Ellsberg, D. (1961). Risk, ambiguity and Savage axioms. *Q. J. Economics* 75, 643–679.

Epstein, S. (1994). Integration of the cognitive and psychodynamic unconscious. *Am. Psychol.* 49, 709–716.

Erev, I. and Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Rev.* 112, 912–931.

Ericsson, K.A. and Simon, H.A. (1993). *Protocol Analysis: Verbal Reports as Data*, revised edition. Cambridge, MA: MIT Press.

Figner, B., Grinband, J., Bayer, H., *et al*. (2007). Neural correlates of risk and return in risky decision making. Working Paper, Center for the Decision Sciences (CDS), Columbia University.

Figner, B., Mackinlay, R. J., Wilkening, F., and Weber, E.U. (2007). Hot and cold cognition in risky decision making: accounting for age and gender differences in risk taking. Working Paper, Center for the Decision Sciences, Columbia University.

Ford, J.K., Schmitt, N., Schechtman, S.L., and Hults, B.M. (1989). Process tracing methods: contributions, problems, and neglected research questions. *Org. Behav. Hum. Dec. Proc.* 43, 75–117.

Fox, C.R. and Tversky, A. (1995). Ambiguity aversion and comparative ignorance. *Q. J. Economics* 110, 879–895.

Gabaix, X., Laibson, D., Moloche, G., and Weinberg, S. (2006). Information acquisition: experimental analysis of a boundedly rational model. *Am. Econ. Rev.* 96, 1043–1068.

Gaechter, S., Johnson, E.J., and Herrmann, A. (2007) Individual-level loss aversion in riskless and risky choices. IZA Discussion Paper No. 2961, available at SSRN: http://ssrn.com/abstract=1010597.

Glimcher, P.W. (2003). *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics*. Boston, MA: MIT Press.

Hanoch, Y., Johnson, J.G., and Wilke, A. (2006). Domain specificity in experimental measures and participant recruitment: an application to risk-taking behavior. *Psychological Sci.* 17, 300–304.

Hertwig, R., Barron, G., Weber, E.U., and Erev, I. (2004). Decisions from experience and the effect of rare events. *Psychological Sci.* 15, 534–539.

Hogarth, R.M. and Kunreuther, H. (1989). Risk, ambiguity and insurance. *J. Risk Uncertainty* 2, 5–35.

Holt, C.A. and Laury, S.K. (2001). Risk aversion and incentive effects. *Am. Econ. Rev.* 92, 1644–1655.

Huettel, S.A. (2006). Behavioral, but not reward, risk modulates activation of prefrontal, parietal, and insular cortices. *Cogn. Affect. Behav. Neurosci.* 6(2), 142–152.

Huettel, S.A., Stowe, C.J., Gordon, E.M. *et al*. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron* 49, 765–775.

Hsu, M., Bhatt, M., Adolphs, R. *et al*. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310, 1680–1683.

Jia, J. and Dyer, J.S. (1997). A standard measure of risk and risk-value models. *Eur. J. Operational Res.* 103, 531–546.

Johnson, E.J. and Goldstein, D. (2003). Do defaults save lives? *Science* 302, 1338–1339.

Johnson, E.J., Schulte-Mecklenbeck, M., and Willemsem, M. (2008). Process models deserve process data: a comment on Brandstaetter, Hertwig and Gigerenzer (2006). *Psychological Rev.* 115, 272–273.

Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *Am. Psychologist* 58, 697–704.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–292.

Keller, L.R. (1985). An empirical investigation of relative risk aversion. *IEEE Trans. Syst. Man Cybern.* SMC15, 475–482.

Keysers, C., Cohen, J., Donald, M. *et al.* (2008). Explicit and implicit strategies in decision making. In: C. Engel and W. Singer (eds), *Better than Conscious? Implications for Performance and Institutional Analysis*. Strungman Forum Report. Cambridge, MA: MIT Press. 225–258.

Knight, F.H. (1921/1964). *Risk, Uncertainty, and Profit*. New York, NY: Sentry Press.

Laciana, C.E., Weber, E.U., Bert, F. et al. (2007). Objective functions in agricultural decision-making: a comparison of the effects of expected utility, regret-adjusted expected utility, and prospect theory maximization. Working Paper, Center for Research on Environmental Decisions (CRED), Columbia University.

Landman, J. (1993). *Regret: The Persistence of the Possible*. Oxford: Oxford University Press.

Lejuez, C.W., Read, J.P., Kahler, C.W. *et al*. (2002). Evaluation of a behavioral measure of risk-taking: the Balloon Analogue Risk Task (BART). *J. Exp. Psychol. Appl.* 8, 75–84.

Lempert, R.J., Nakicenovic, N., Sarewitz, D., and Schlesinger, M.E. (2004). Characterizing climate-change uncertainties for decision-makers. *Climatic Change* 65, 1–9.

Levy, H. and Markowitz, H. (1979). Approximating expected utility by a function of mean and variance. *Am. Econ. Rev.* 9, 308–317.

Lichtenstein, S. and Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *J. Exp. Psychol.* 89, 46–55.

Loewenstein, G.F., Weber, E.U., Hsee, C.K., and Welch, E. (2001). Risk as feelings. *Psychological Bull.* 127, 267–286.

Loomes, G. and Sugden, R. (1982). Regret theory: an alternative theory of rational choice under uncertainty. *Economic J.* 92, 805–824.

Lopes, L.L. (1987). Between hope and fear: the psychology of risk. *Adv. Exp. Social Psychol.* 20, 255–295.

Luce, R.D. and Weber, E.U. (1986). An axiomatic theory of conjoint, expected risk. *J. Math. Psychol.* 30, 188–205.

MacCrimmon, K.R. and Wehrung, D.A. (1986). *Taking Risks: The Management of Uncertainty*. New York, NY: Free Press.

Machina, M.J. (1987). Choice under uncertainty: problems solved and unsolved. *J. Econ. Persp.* 12, 121–154.

Markman, K.D., Gavanski, I., Sherman, S.J., and McMullen, M.N. (1993). The mental simulation of better and worse possible worlds. *J. Exp. Social Psychol.* 29, 87–109.

Markowitz, H.M. (1952). Portfolio selection. *J. Finance* 7, 77–91.

Markowitz, H.M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. New York, NY: John Wiley & Sons.

Marsh, B. and Kacelnik, A. (2002). Framing effects and risky decisions in starlings. *PNAS* 99, 3352–3355.

McFadden, D. (1999). Rationality for economists? *J. Risk Uncertainty* 19, 73–105.

Mellers, B.A., Ordonez, L.D., and Birnbaum, M.H. (1992). A change-of-process theory for contextual effects and preference reversals in risky decision making. Special Issue: Utility measurement. *Org. Behav. Hum. Dec. Proc.* 52, 331–369.

Nosic, A. and Weber, M. (2007). Determinants of risk taking behavior: the role of risk attitudes, risk perception, and beliefs. Working Paper, University of Mannheim.

Payne, J.W., Bettman, J.R., and Johnson, E.J. (1988). Adaptive strategy selection in decision-making. *J. Exp. Psychol. Learning Memory Cogn.* 14, 534–552.

Payne, J.W., Bettman, J.R., and Johnson, E.J. (1993). *The Adaptive Decision-Maker*. Cambridge: Cambridge University Press.

Piaget, J. (1964). *Six Psychological Studies*. New York, NY: Vintage.

Pinker, S. (1994). *The Language Instinct*. New York, NY: William Morrow.

Pratt, J.W. (1964). Risk aversion in the small and in the large. *Econometrica* 32, 122–136.

Preuschoff, K., Bossaerts, P., and Quartz, S. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51, 381–390.

Rabin, M. (1998). Psychology and economics. *J. Econ. Lit.* 36, 11–46.

Rettinger, D.A. and Hastie, R. (2002). Content effects on decision making. *Org. Behav. Hum. Dec. Proc.* 85, 336–359.

Russo, J.E. and Dosher, B.A. (1983). Strategies for multiattribute binary choice. *J. Exp. Psychol. Learning Memory Cogn.* 9, 676–696.

Samuelson, W.F. and Zeckhauser, R.J. (1988). Status quo bias in decision making. *J. Risk Uncertainty* 1, 7–59.

Sarin, R.K. and Weber, M. (1993). Risk–value models. *Eur. J. Operations Res.* 70, 135–149.

Schkade, D.A. and Johnson, E.J. (1989). Cognitive–processes in preference reversals. *Org. Behav. Hum. Dec. Proc.* 44(2), 203–231.

Sharpe, and William, F. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *J. Finance* 19, 425–442.

Sloman, S.A. (1996). The empirical case for two systems of reasoning. *Psychological Bull.* 1, 3–22.

Slovic, P., Finucane, M., Peters, E., and MacGregor, D.G. (2002). The affect heuristic. In: T. Gilovich, D. Griffin, and D. Kahneman (eds), *Heuristics and Biases*. Cambridge: Cambridge University Press, pp. 397–420.

Stanovich, K.E. and West, R.F. (2000). Individual differences in reasoning: implications for the rationality debate. *Behav. Brain Sci.* 23, 645–665.

Steinberg, L. (2007). Risk taking in adolescence: new perspectives from brain and behavioral science. *Curr. Dir. Psychol. Sci.* 16, 55–59.

Tom, S., Fox, C.R., Trepel, C., and Poldrack, R.A. (2007). The neural basis of loss-aversion in decision-making under risk. *Science* 315, 515–518.

Thaler, R.H. (1980). Toward a positive theory of consumer choice. *J. Econ. Behav. Org.* 1, 39–60.

Thaler, R.H. and Johnson, E.J. (1990). Gambling with the house money and trying to break even: the effects of prior outcomes in risky choice. *Management Sci.* 36, 643–660.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory, cumulative representation of uncertainty. *J. Risk Uncertainty* 5, 297–323.

von Neumann, J. and Morgenstern, O. (1944/1947). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

von Winterfeldt, D. and Edwards, W. (1986). *Decision Analysis and Behavioral Research*. Cambridge: Cambridge University Press.

Wallsten, T.S., Pleskac, T.J., and Lejuez, C.W. (2005). Modeling behavior in a clinically diagnostic sequential risk-taking task. *Psychological Rev.* 112, 862–880.

Weber, E.U. (1994). From subjective probabilities to decision weights: the effect of asymmetric loss functions on the evaluation of uncertain outcomes and events. *Psychological Bull.* 115, 228–242.

Weber, E.U. (2001a). Decision and choice: risk, empirical studies. In: N.J. Smelser and P.B. Baltes (eds), *International Encyclopedia of the Social and Behavioral Sciences*. Oxford: Elsevier, pp. 13347–13351.

Weber, E.U. (2001b). Personality and risk taking. In: N.J. Smelser and P.B. Baltes (eds), *International Encyclopedia of the Social and Behavioral Sciences*. Oxford: Elsevier, pp. 11274–11276.

Weber, E.U. (2004). Perception matters: psychophysics for economists. In: J. Carrillo and I. Brocas (eds), *Psychology and Economics*. Oxford: Oxford University Press, pp. 165–176.

Weber, E.U. (2006). Experience-based and description-based perceptions of long-term risk: why global warming does not scare us (yet). *Climatic Change* 70, 103–120.

Weber, E.U. and Hsee, C.K. (1998). Cross-cultural differences in risk perception but cross-cultural similarities in attitudes towards risk. *Management Sci.* 44, 1205–1217.

Weber, E.U. and Kirsner, B. (1997). Reasons for rank-dependent utility evaluation. *J. Risk Uncertainty* 14, 41–61.

Weber, E.U. and Milliman, R. (1997). Perceived risk attitudes: relating risk perception to risky choice. *Management Sci.* 43, 122–143.

Weber, E.U., Blais, A.R., and Betz, N. (2002). A domain-specific risk attitude scale: measuring risk perceptions and risk behaviors. *J. Behav. Decision Making* 15, 263–290.

Weber, E.U., Shafir, S., and Blais, A.R. (2004). Predicting risk-sensitivity in humans and lower animals: risk as variance or coefficient of variation. *Psychological Rev.* 111, 430–445.

Weber, E.U., Siebenmorgen, N., and Weber, M. (2005). Communicating asset risk: how name recognition and the format of historic volatility information affect risk perception and investment decisions. *Risk Analysis* 25, 597–609.

Weber, M. and Zuchel, H. (2005). How do prior outcomes affect risk attitude? Comparing escalation of commitment and the house money effect. *Decision Analysis* 2, 30–43.

Zentall, T.R., Jr, Galef, B.G., and Zentall, T.R. (eds) (1988). *Social Learning: Psychological and Biological Perspectives*. Hillsdale, NJ: Lawrence Erlbaum.

Zuckerman, M. (1979). *Sensation Seeking: Beyond the Optimal Level of Arousal*. Hillsdale, NJ: Lawrence Erlbaum.

Zuckerman, M., Simons, R.F., and Como, P.G. (1988). Sensation seeking and stimulus intensity as modulators of cortical, cardiovascular, and electrodermal response: a cross-modality study. *Pers. Indiv. Diff.* 9, 361–372.

CHAPTER

11

Prospect Theory and the Brain

Craig R. Fox and Russell A. Poldrack

OUTLINE

INTRODUCTION TO PROSPECT THEORY

Whether we like it or not, we face risk every day of our lives. From selecting a route home from work to selecting a mate, we rarely know in advance and with certainty what the outcome of our decisions will be. Thus, we are forced to make tradeoffs between the attractiveness (or unattractiveness) of potential outcomes and their likelihood of occurrence.

The lay conception of "risk" is associated with hazards that fill one with dread or are poorly understood (Slovic, 1987). Managers tend to see risk not as a gamble but as a "challenge to be overcome," and see risk as increasing with the magnitude of potential losses (e.g., March and Shapira, 1987). Decision theorists, in contrast, view risk as increasing with variance in the probability distribution of possible outcomes, regardless of whether a potential loss is involved. For example, a prospect that offers a 50–50 chance of paying $100 or nothing is more risky than a prospect that offers $50 for sure – even though the "risky" prospect entails no possibility of losing money.

Since Knight (1921), economists have distinguished decisions under *risk* from decisions under *uncertainty*. In decisions under risk, the decision maker knows with precision the probability distribution of possible outcomes, as when betting on the flip of a coin or entering a lottery with a known number of tickets.

In decisions under uncertainty, the decision maker is not provided such information but must assess the probabilities of potential outcomes with some degree of vagueness, as when betting on a victory by the home team or investing in the stock market.

In this chapter, we explore behavioral and neuroeconomic perspectives on decisions under risk. For simplicity we will confine most of our attention to how people evaluate simple prospects with a single non-zero outcome that occurs with known probability (e.g., a 50–50 chance of winning \$100 or nothing), though we will also mention extensions to multiple outcomes and to vague or unknown probabilities.

In the remainder of this section we provide a brief overview of economic models of decision making under risk, culminating in prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), the most influential descriptive account that has emerged to date. In subsequent sections, we provide an overview of various parameterizations of prospect theory's functions, and review methods for eliciting them. We then take stock of the early neuroeconomic studies of prospect theory, before providing some suggested directions for future research.

Historical Context

The origin of decision theory is traditionally traced to a correspondence between Pascal and Fermat in 1654 that laid the mathematical foundation of probability theory. Theorists asserted that decision makers ought to choose the option that offers the highest expected value (EV). Consider a prospect (x, p) that offers \$$x$ with probability p (and nothing otherwise):

$$\mathrm{EV} = px. \tag{11.1}$$

A decision maker is said to be "risk neutral" if he is indifferent between a gamble and its expected value; he is said to be "risk averse" if he prefers a sure payment to a risky prospect of equal or higher expected value; he is said to be "risk seeking" if he prefers a risky prospect to a sure payment of equal or higher expected value. Thus, expected value maximization assumes a neutral attitude toward risk. For instance, a decision maker who employs this rule will prefer receiving \$100 if a fair coin lands heads (and nothing otherwise) to a sure payment of \$49, because the expected value of the gamble (\$50 = .5 × \$100) is higher than the value of the sure thing (\$49).

Expected value maximization is problematic because it does not allow decision makers to exhibit risk aversion – it cannot explain, for example, why a person would prefer a sure \$49 over a 50–50 chance of receiving \$100 or nothing, or why anyone would purchase insurance. Swiss mathematician Daniel Bernoulli (1738) advanced a solution to this problem when he asserted that people do not evaluate options by their objective value but rather by their utility or "moral value." Bernoulli observed that a particular amount of money (say, \$1000) is valued more when a person is poor (wealth level W_1) than when he is wealthy (W_2) and therefore marginal utility decreases (from U_1 to U_2) as wealth increases (see Figure 11.1a). This gives rise to a utility function that is concave over states of wealth. In Bernoulli's model, decision makers choose the option with highest expected utility (EU):

$$EU = pu(x) \tag{11.2}$$

where $u(x)$ represents the utility of obtaining outcome x. For example, a concave utility function ($u''(x) < 0$) implies that the utility gained by receiving

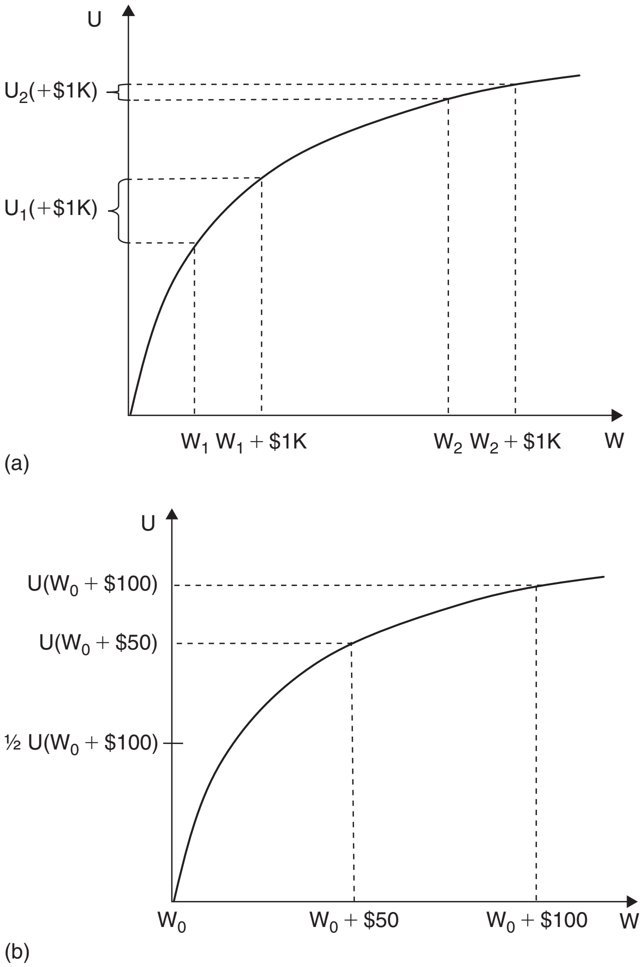


FIGURE 11.1 (a) A representative utility function over states of wealth illustrating the notion of diminishing marginal utility. (b) A representative utility function over states of wealth illustrating risk aversion for gains at an initial state of wealth W_0.

\$50 is more than half the utility gained by receiving \$100, and therefore a decision maker with such a utility function should prefer \$50 for sure to a .5 probability of receiving \$100 (see Figure 11.1b)

Axiomatization of Expected Utility

Expected utility became a central component of economic theory when von Neumann and Morgenstern (1947) articulated a set of axioms that are both necessary and sufficient for representing a decision-maker's choices by the maximization of expected utility (see also Jensen, 1967). Consider chance lotteries L_1 and L_2 that are known probability distributions over outcomes. For instance, L_1 might offer a .5 chance of \$100 and a .5 chance of 0; L_2 might offer \$30 for sure. Consider also a binary preference relation $\succeq$ over the set of all possible lotteries L; thus $L_1 \succeq L_2$ is interpreted as "L_1 is preferred or equivalent to L_2." Now consider the following axioms:

1 **Completeness**: People have preferences over all lotteries. Formally, for any two lotteries L_1 and L_2 in L, either $L_1 \succeq L_2$, $L_2 \succeq L_1$, or both.

2 **Transitivity**: People rank lotteries in a consistent manner. Formally, for any three lotteries L_1, L_2, and L_3, if $L_1 \succeq L_2$, and $L_2 \succeq L_3$, then $L_1 \succeq L_3$.

3 **Continuity**: For any three lotteries, some mixture of the best and worst lotteries is preferred to the intermediate lottery and *vice versa*. Formally, for any three lotteries $L_1 \succeq L_2 \succeq L_3$ there exist α, $\beta \in (0,1)$ such that $\alpha L_1 + (1 - \alpha) L_3 \succeq L_2$, and $L_2 \succeq \beta L_1 + (1 - \beta) L_3$.

4 **Substitution** (a.k.a. "independence"): If a person prefers one lottery to another, then this preference should not be affected by a mixture of both lotteries with a common third lottery. Formally, for any L_1, L_2, and L_3, and any $\alpha \in (0, 1)$, $L_1 \succeq L_2$ if and only if $\alpha L_1 + (1 - \alpha) L_3 \succeq \alpha L_2 + (1 - \alpha) L_3$.

Von Neumann and Morgenstern proved that these axioms are both necessary and sufficient to represent a decision-maker's decisions by the maximization of expected utility. That is,

$$L_1 \succeq L_2 \text{ if and only if } \sum_{i=1}^{n} p_i^1 u(x_i^1) \geq \sum_{j=1}^{m} p_j^2 u(x_j^2),$$

where superscripts indicate corresponding lottery numbers.

The completeness and transitivity axioms establish that decision makers can (weakly) order their preferences, which is necessary for using a unidimensional scale. The continuity axiom is necessary to establish a continuous tradeoff between probability and outcomes. The substitution axiom is necessary to establish that utilities of outcomes are weighted by their respective probabilities.

A more general formulation of expected utility theory that extended the model from risk to uncertainty (Savage, 1954) relies on a related axiom known as the *sure-thing principle*: If two options yield the same consequence when a particular event occurs, then a person's preferences among those options should not depend on the particular consequence (i.e., the "sure thing") or the particular event that they have in common. To illustrate, consider a game show in which a coin is flipped to determine where a person will be sent on vacation. Suppose the contestant would rather to go to Atlanta if the coin lands heads and Chicago if it lands tails (a, H; c, T) than go to Boston if the coin lands heads and Chicago if it lands tails (b, H; c, T). If this is the case, he should also prefer to go to Atlanta if the coin lands heads and Detroit (or any other city for that matter) if the coin lands tails (a, H; d, T), to Boston if it lands heads and Detroit if it lands tails (b, H; d, T).

Violations of Substitution and the Sure thing Principle

It was not long before the descriptive validity of expected utility theory and its axioms were called into question. One of the most powerful challenge has come to be known as the "Allais paradox" (Allais, 1953; Allais and Hagen, 1979). The following version was presented by Kahneman and Tversky (1979)[1].

Decision 1: Choose between (A) an 80% chance of \$4000; (B) \$3000 for sure.
Decision 2: Choose between (C) a 20% chance of \$4000; (D) a 25% chance of \$3000.

Most respondents chose (B) over (A) in the first decision and (C) over (D) in the second decision, which violates the substitution axiom. To see why, note that $C = 1/4\ A$ and $D = 1/4\ B$ (with a $3/4$ chance of receiving 0 in both cases) so that according to the substitution axiom a decision maker should prefer C over D if and only if he prefers A to B. This systematic violation of substitution is known as the "common ratio effect."

A related demonstration from Allais was adapted by Kahneman and Tversky (1979) as follows:

Decision 3: Choose between (E) a 33% chance of \$2500, a 66% chance of \$2400, and a 1% chance of nothing; (F) \$2400 for sure.
Decision 4: Choose between (G) a 33% chance of \$2500; (H) a 34% chance of \$2400.

---

[1] Kahneman & Tversky's version was originally denominated in Israeli Pounds.

TABLE 11.1 The Allais common consequence effect represented using a lottery with numbered tickets

	Ticket numbers		
Option	**1–33**	**34**	**35–100**
E	2500	0	2400
F	2400	2400	2400
G	2500	0	0
H	2400	2400	0

TABLE 11.2 The fourfold pattern of risk attitudes (a); risk aversion for mixed (gain–loss) gambles (b) (both adapted from Tversky and Kahneman, 1992)

(a) $C(x, p)$ is the median certainty equivalent of the prospect that pays $\$x$ with probability p

	Gains	**Losses**
Low probability	C ($100, .05) = $14 *Risk seeking*	C (−$100, .05) = − $8 *Risk aversion*
High probability	C ($100, .95) = $78 *Risk aversion*	C (−$100, .95) = − $84 *Risk-seeking*

(b) Median gain amounts for which participants found 50–50 mixed gambles equally attractive to receiving nothing, listed fixed by loss amount

Gain	**Loss**	**Ratio**
61	25	2.44
101	50	2.02
202	100	2.02
280	150	1.87

In this case most people prefer option (F) to option (E) in Decision 3, but they prefer option (G) to option (H) in Decision 4, which violates the sure-thing principle. To see why, consider options (E) through (H) as being payment schemes attached to different lottery tickets that are numbered consecutively from 1 to 100 (see Table 11.1). Note that one can transform options (E) and (F) into options (G) and (H), respectively, merely by replacing the common consequence (receive $2400 if the ticket drawn is 35–100) with a new common consequence (receive $0 if the ticket drawn is 35–100). Thus, according to the sure-thing principle, a person should favor option (G) over option (H) if and only if he prefers option (E) to option (F), and the dominant pattern of preferences violates this axiom. This violation of the sure-thing principle is known as the "common consequence effect."

Both the common ratio effect and common consequence effect resonate with the notion that people are more sensitive to differences in probability near impossibility and certainty than in the intermediate range of the probability scale. Thus, people typically explain their choice in Decision (1) as a preference for certainty over a slightly smaller prize that entails a possibility of receiving nothing; meanwhile, they explain their choice in Decision (2) as a preference for a higher possible prize given that the between a difference in probability of .20 and .25 is not very large. Likewise, people explain their choice in Decision (3) as a preference for certainty over a possibility of receiving nothing; meanwhile, they explain their choice in Decision (2) as a preference for a higher possible prize given that the difference between a probability of .33 and .34 seems trivial.

The Fourfold Pattern of Risk Attitudes

The Allais paradox is arguably the starkest and most celebrated violation of expected utility theory. In the years since it was articulated, numerous studies of decision under risk have shown that people often violate the principle of risk aversion that underlies much economic analysis. Table 11.2 illustrates a common pattern of risk aversion and risk seeking exhibited by participants in studies of Tversky and Kahneman (1992). Let $C(x, p)$ be the *certainty equivalent* of the prospect (x, p) that offers to pay $\$x$ with probability p (i.e., the sure payment that is deemed equally attractive to the risky prospect). The upper left-hand entry in Table 11.2 shows that the median participant was indifferent between receiving $14 for sure and a 5% chance of gaining $100. Because the expected value of the prospect is only $5, this observation reflects risk seeking behavior.

Table 11.2a reveals a fourfold pattern of risk attitudes: risk seeking for low-probability gains and high-probability losses, coupled with risk aversion for high-probability gains and low-probability losses. Choices consistent with this fourfold pattern have been observed in several studies (Fishburn and Kochenberger, 1979; Kahneman and Tversky, 1979; Hershey and Schoemaker, 1980; Payne *et al.*, 1981). Risk seeking for low-probability gains may contribute to the attraction of gambling, whereas risk aversion for low-probability losses may contribute to the attraction of insurance. Risk aversion for high-probability gains may contribute to the preference for certainty, as in the Allais (1953) problem, whereas risk seeking for high-probability losses is consistent with the common tendency to undertake risk to avoid facing a sure loss.

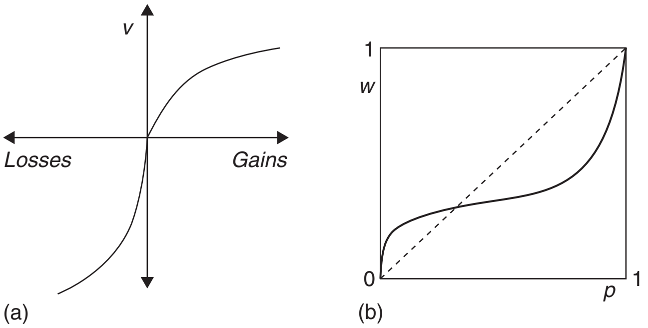


FIGURE 11.2 Representative value and weighting functions from prospect theory. (a) A hypothetical prospect theory value function illustrating concavity for gains, convexity for losses, and a steeper loss than gain limb. (b) A hypothetical prospect theory weighting function illustrating its characteristics inverse-S shape, the tendency to overweight low probabilities and underweight moderate to large probabilities, and the tendency for weights of complementary probabilities to sum to less than 1.

Prospect Theory

The Allais paradox and the fourfold pattern of risk attitudes are accommodated neatly by prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), the leading behavioral model of decision making under risk, and the major work for which psychologist Daniel Kahneman was awarded the 2002 Nobel Prize in economics.

According to prospect theory, the value V of a simple prospect that pays \$$x$ with probability p (and nothing otherwise) is given by:

$$V(x, p) = w(p)\, v(x) \tag{11.3}$$

where v measures the subjective value of the consequence x, and w measures the impact of probability p on the attractiveness of the prospect (see Figure 11.2).

Value Function

Prospect theory replaces the utility function $u(\cdot)$ over states of wealth with a value function $v(\cdot)$ over gains and losses relative to a reference point, with $v(0) = 0$. According to prospect theory, the value function $v(\cdot)$ exhibits the psychophysics of diminishing sensitivity. That is, the marginal impact of a change in value diminishes with the distance from a relevant reference point. For monetary outcomes, the *status quo* generally serves as the reference point distinguishing losses from gains, so that the function is concave for gains and convex for losses (see Figure 11.2a). Concavity for gains contributes to risk aversion for gains, as with the standard utility function (Figure 11.1). Convexity for losses, on the other hand, contributes to risk seeking for losses. For instance, the disvalue of losing \$50 is more than half the disvalue of losing \$100, which will contribute to a preference for the gamble over the sure loss. This tendency to be risk averse for moderate-probability gains and risk seeking for moderate-probability losses may contribute to the "disposition effect," in which investors have a greater tendency to sell stocks in their portfolios that have risen rather than fallen since purchase (Odean, 1998; but see also Barberis and Xiong, 2006).

The prospect theory value function is steeper for losses than gains – a property known as *loss aversion*. People typically require more compensation to give up a possession than they would have been willing to pay to obtain it in the first place (see, for example, Kahneman *et al*., 1990). In the context of decision under risk, loss aversion gives rise to risk aversion for mixed (gain–loss) gambles so that, for example, people typically reject a gamble that offers a .5 chance of gaining \$100 and a .5 chance of losing \$100, and require at least twice as much "upside" as "downside" to accept such gambles (see Table 11.2b). In fact, Rabin (2000) showed that a concave utility function over states of wealth cannot explain the normal range of risk aversion for mixed gambles, because this implies that a decision maker who is mildly risk averse for small-stakes gambles over a range of states of wealth must be unreasonably risk averse for large-stakes gambles. This tendency to be risk averse for mixed prospects has been used by Benartzi and Thaler (1995) to explain why investors require a large premium to invest in stocks rather than bonds (the "equity premium puzzle"): because of the higher volatility of stocks than bonds, investors who frequently check their returns are more likely to experience a loss in nominal value of their portfolios if they are invested in stocks than bonds (see also Barberis *et al*., 2001).

It is important to note that loss aversion, which gives rise to risk aversion for mixed (gain–loss) prospects (e.g., most people reject a 50–50 chance to gain \$100 or lose \$100) should be distinguished from convexity of the value function for losses, which gives rise to risk-seeking for pure loss prospects (e.g., most people prefer a 50–50 chance to lose \$100 or nothing, to losing \$50 for sure).

Weighting Function

In prospect theory, the value of an outcome is weighted not by its probability but instead by a decision weight, $w(\cdot)$, that represents the impact of the relevant probability on the valuation of the prospect (see equation 11.3). Decision weights are normalized so that $w(0) = 0$ and $w(1) = 1$. Note that w need not be interpreted as a measure of subjective belief – a person may believe that the probability of a fair coin landing heads is one-half, but afford this event a weight of less than one-half in the evaluation of a prospect.

Just as the value function captures diminishing sensitivity to changes in the number of dollars gained or lost, the weighting function captures diminishing sensitivity to changes in probability. For probability, there are two natural reference points: impossibility and certainty. Hence, diminishing sensitivity implies an inverse-S shaped weighting function that is concave near zero and convex near one, as depicted in Figure 11.2b. It can help explain the fourfold pattern of risk attitudes (Table 11.2a), because moderate to high probabilities are underweighted (which reinforces the pattern of risk aversion for gains and risk seeking for losses implied by the shape of the value function) and low probabilities are overweighted (which reverses the pattern implied by the value function and leads to risk seeking for gains and risk aversion for losses).

To appreciate the intuition underlying how the value- and weighting-functions contribute to the fourfold pattern, refer to Figure 11.2. Informally, the reason that most participants in Tversky and Kahneman's (1992) sample would rather have a .95 chance of $100 than $77 for sure is partly because they find receiving $77 nearly as appealing as receiving $100 (i.e., the slope of the value function decreases with dollars gained), and partly because a .95 chance "feels" like a lot less than a certainty (i.e., the slope of the weighting function is high near one). Likewise, most participants would rather face a .95 chance of losing $100 than pay $85 for sure is partly because paying $85 is almost as painful as paying $100, and partly because a .95 chance feels like it is much less than certain. On the other hand, the reason that most participants would rather have a .05 chance of $100 than $13 for sure is that a .05 chance "feels" like much more than no chance at all (i.e., the slope of the weighting function is steep near zero) – in fact it "feels" like more than its objective probability, and this distortion is more pronounced than the feeling that receiving $13 is more than 13% as attractive as receiving $100. Likewise, the reason most participants would rather lose $7 for sure than face a .05 chance of losing $100 is that the .05 chance of losing money looms larger than its respective probability, and this effect is more pronounced than the feeling that receiving $7 is more than 7% as attractive as receiving $100.

The inverse-S shaped weighting function also explains the Allais paradox because the ratio of weights of probabilities .8 and 1 is smaller than the ratio of weights of probabilities .20 and .25 (so that the difference between a .80 chance of a prize and a certainty of a prize in Decision 1 looms larger than the difference between a .20 and .25 chance of a prize in Decision 2); similarly, the difference in the weights of probabilities .99 and 1 is larger than the difference in the weights of probabilities .33 and .34 (so that the difference between a .99 chance and a certainty of receiving a prize in Decision 3 looms larger than the difference between a .33 chance and a .34 chance in Decision 4). This inverse S-shaped weighting function seems to be consistent with a range of empirical findings in laboratory studies (e.g., Camerer and Ho, 1994; Tversky and Fox, 1995; Wu and Gonzalez, 1996, 1998; Gonzalez and Wu, 1999; Wakker, 2001). Overweighting of low-probability gains can help explain why the attraction of lotteries tends to increase as the top prize increases even as the chances of winning decreases correspondingly (Cook and Clotfelter, 1993) and the attraction to long-shot bets over favorites in horse races. Overweighting of low-probability losses can also explain the attractiveness of insurance (Wakker *et al*., 1997).

In sum, prospect theory explains attitudes toward risk via distortions in shape of the value and weighting functions. The data of Tversky and Kahneman (1992) suggest that the fourfold pattern of risk attitudes for simple prospects that offer a gain or a loss with low or high probability (Table 11.2a) is driven primarily by curvature of the weighting function, because the value function is not especially curved for the typical participant in those studies. Pronounced risk aversion for mixed prospects that offer an equal probability of a gain or loss (Table 11.2b) is driven almost entirely by loss aversion, because the curvature of the value function is typically similar for losses versus gains and decision weights are similar for gain versus loss components.

Framing and Editing

Expected utility theory and most normative models of decision making under risk assume *description invariance*: preferences among prospects should not be affected by how they are described. Decision makers should act as if they are assessing the impact of options on final states of wealth. Prospect theory, in contrast, explicitly acknowledges that choices are influenced by how prospects are cognitively represented in terms of losses and gains and their associated probabilities. There are two important manifestations of this principle.

First, this representation can be systematically influenced by the way in which options are described or "framed." Recall that the value function is applied to a reference point that distinguishes between losses and gains. A common default reference point is the *status quo*. However, by varying the description of options one can influence how they are perceived. For instance, decisions concerning medical treatments can differ depending on whether possible outcomes are described in terms of survival versus mortality rates (McNeil *et al.*, 1982); recall that people tend to be risk averse for moderate probability gains and risk seeking for moderate probability losses. Likewise, the weighting function is applied to probabilities of risky outcomes that a decision maker happens to identify. The description of gambles can influence whether probabilities are integrated or segregated, and therefore affect the decisions that people make (Tversky and Kahneman, 1986). For instance, people were more likely to favor a .25 chance of $32 over a .20 chance of $40 when this choice was described as a two-stage game in which there was a .25 chance of obtaining a choice between $32 for sure or a .80 chance of $40 (that is, the $32 outcome was more attractive when it was framed as a certainty). People may endogenously frame prospects in ways that are not apparent to observers, adopting aspirations as reference points (Heath *et al.*, 1999) or persisting in the adoption of old reference points, viewing recent winnings as "house money" (Thaler and Johnson, 1990).

Second, people may mentally transform or "edit" the description of prospects they have been presented. The original formulation of prospect theory (Kahneman and Tversky, 1979) suggested that decision makers edit prospects in forming their subjective representation. Consider prospects of the form ($\$x_1$, p_1; $\$x_2$, p_2; $\$x_3$, p_3) that offer $\$x_i$ with (disjoint) probability p_i (and nothing otherwise). In particular, decision makers are assumed to engage in the following mental transformations:

1. *Combination*. Decision makers tend to simplify prospects by combining common outcomes – for example, a prospect that offers ($10, .1; $10, .1) would be naturally represented as ($10, .2).
2. *Segregation*. Decision makers tend to segregate sure outcomes from the representation of a prospect – for instance, a prospect that offers ($20, .5; $30, .5) would be naturally represented as $20 for sure plus a ($10, .5).
3. *Cancellation*. Decision makers tend to cancel shared components of options that are offered together – for example, a choice between ($10, .1; $50, .1) or ($10, .1; $20, .2) would be naturally represented as a choice between a ($50, .1) or ($20, .2).
4. *Rounding*. Decision makers tend to simplify prospects by rounding uneven numbers or discarding extremely unlikely outcomes – for example, ($99, .51; $5, .0001) might be naturally represented as ($100, .5).
5. *Transparent dominance*. Decision makers tend to reject options without further evaluation if they are obviously dominated by other options – for instance, given a choice between ($18, .1; $19, .1; $20, .1) or ($20, .3), most people would naturally reject the first option because it is stochastically dominated by the second.

Applications to Riskless Choice

Although prospect theory was originally developed as an account of decision making under risk, many manifestations of this model in riskless choice have been identified in the literature.

Loss Aversion

Loss aversion implies that preferences among consumption goods will systematically vary with one's reference point (Kahneman and Tversky, 1991; see also Bateman *et al.*, 1997), which has several manifestations. First, the minimum amount of money a person is willing to accept (WTA) to part with an object generally exceeds the minimum amount of money that he is willing to pay (WTP) to obtain the same object. This pattern, robust in laboratory studies using student populations and ordinary consumer goods, is even more pronounced for non-market goods, non-student populations, and when incentives are included to encourage non-strategic responses (Horowitz and McConnell, 2002).

Likewise, people tend to value objects more highly after they come to feel that they own them – a phenomenon known as the *endowment effect* (Thaler, 1980). For instance, in one well-known study Kahneman *et al.* (1990) presented a coffee mug with a university logo to one group of participants ("sellers") and told them the mug was theirs to keep, then asked these participants whether they would sell the mug back to them at various prices. A second group of participants ("choosers") were told that they could have the option of receiving an identical mug or an amount of money, and asked which they preferred at various prices. Although both groups were placed in strategically identical situations (walk away with a mug or money), the sellers, who presumably framed the choice as a *loss* of a mug against a compensating gain of money, quoted a median price of $7.12, whereas the buyers, who presumably framed the choice as a *gain* of a mug against a gain of money, quoted a median price of $3.12.

Loss aversion is thought to contribute to the inertial tendency to stick with *status quo* options (Samuelson and Zeckhauser, 1988) and the reluctance to trade. For instance, in one study Knetsch (1989) provided students with a choice between a university mug and a bar of Swiss chocolate, and found that they had no significant preference for one over the other. However, when some students were assigned at random to receive the mug and given an opportunity to trade for the chocolate, 89% retained the mug; when other students were assigned at random to receive the chocolate and given an opportunity to trade for the mug, only 10% opted for the mug.

Loss aversion has been invoked to help explain a number of anomalous patterns in field data. Notably, loss aversion can partly account for the powerful attraction of defaults on behavior – for instance, why organ donation rates are much higher for European countries with an "opt-out" policy than those with an "opt-in" policy (Johnson and Goldstein, 2003), the tendency of consumer demand to be more sensitive to price increases than decreases (Hardie *et al*., 1993), and the tendency for taxi drivers to quit after they have met their daily income targets, even on busy days during which their hourly wages are higher (Camerer *et al*., 1997). In fact, Fehr and Gotte (2007) found a similar pattern among bicycle messengers in which only those who exhibited loss-averse preferences for mixed gambles tended to exert less effort per hour when their wage per completed job increased.

The stronger response to losses than foregone gains also manifests itself in evaluations of fairness. In particular, most people find it unfair for an employer or merchant to raise prices on consumers or to lower wages for workers unless the employer or merchant is defending against losses of their own, and this places a constraint on profit-seeking even when the market clearing price (wage) goes up (down) (Kahneman *et al*., 1986). For instance, people find it more fair to take away a rebate than to impose a price increase on customers; most people think it is unfair for a hardware store to exercise its economic power by raising the price of snow shovels after a snowstorm.

Loss aversion is also evident in riskless choice when consumers face tradeoffs of one product attribute against one another. For instance, Kahneman and Tversky (1991) asked participants to choose between two hypothetical jobs: Job x was characterized as "limited contact with others" and a 20-minute daily commute; Job y was characterized as "moderately sociable" with a 60-minute daily commute. Participants were much more likely to choose Job x if they had been told that their present job was socially isolated with a 10-minute commute than if they had been told it was very social but had an 80-minute commute, consistent with the notion that they are loss averse for relative advantages and disadvantages. Loss aversion when making tradeoffs may partially explain the ubiquity of brand loyalty in the marketplace.

Given the disparate manifestations of loss aversion, it is natural to ask to what extent there is any consistency in a person's degree of loss aversion across these different settings. Johnson *et al*. (2007) approached customers of a car manufacturer and, through a series of simple tasks, determined each customer's coefficient of loss aversion in a risky context, as well as a measure of the endowment effect that compares the minimum amount of money each participant was willing to accept to give up a model car and their maximum willingness to pay to acquire the model car. Remarkably, the Spearman correlation between the risky and riskless measures was .635, suggesting some consistency in the underlying trait of loss aversion.

Curvature of the Value Function

Not only does the difference in steepness of the value function for losses versus gains affect riskless choice, but so does the difference in curvature. Notably, Heath *et al*. (1999) asserted that goals can serve as reference points that inherit properties of the prospect theory value function. For instance, most people believe that a person who has completed 42 sit-ups would be willing to exert more effort to complete one last sit-up if he had set a goal of 40 than if he had set a goal of 30, because the value function is steeper (above the reference point) in the former than in the latter case. Conversely, most people believe that a person who has completed 28 sit-ups would be willing to exert more effort to complete one last sit-up if he had set a goal of 30 than if he had set a goal of 40, because value function is steeper (below the reference point) in the former case than in the latter case.

The cognitive activities that people use to frame and package gains and losses, known as "mental accounting" (Thaler, 1980, 1985, 1999), can influence the way in which riskless outcomes are experienced. In particular, due to the concavity of the value function for gains, people derive more enjoyment when gains are segregated (e.g., it's better to win two lotteries on two separate days); due to the convexity of the value function for losses, people find it less painful when losses are integrated (e.g., it's better to pay a parking ticket the same day I pay my taxes) – but see Linville and Fischer (1991).

Extensions of Prospect Theory

As mentioned earlier, decision theorists distinguish between decisions under risk, in which probabilities are known to the decision maker, and decisions under uncertainty, in which they are not. The original formulation of prospect theory (henceforth OPT; Kahneman and Tversky, 1979) applies to decisions under risk and involving at most two non-zero outcomes. *Cumulative prospect theory* (henceforth CPT; Tversky and Kahneman, 1992;

see also Luce and Fishburn, 1991; Wakker and Tversky, 1993) accommodates decisions under uncertainty and any finite number of possible outcomes. A thorough account of CPT is beyond the scope of this chapter, so we will only sketch out its distinctive features and refer the reader to the original paper for further detail.

Cumulative Prospect Theory

When considering simple chance prospects with at most two non-zero outcomes, two distinctive features of CPT are important.

First, cumulative prospect theory segregates value into gain portions and loss portions, with separate weighting functions for losses and gains (i.e., CPT decision weights are *sign-dependent*)[2].

Second, CPT applies decision weights to cumulative distribution functions rather than single events (i.e., CPT decision weights are *rank-dependent*)[3]. That is, each outcome x is weighted not by its probability but by the cumulated probabilities of obtaining an outcome at least as good as x if the outcome is positive, and at least as bad as x if the outcome is negative.

More formally, consider a chance prospect with two non-zero outcomes $(x, p; y, q)$ that offers \$$x$ with probability p and \$$y$ with probability q (otherwise nothing). Let $w^+(\cdot)$ and $w^-(\cdot)$ be the weighting function for gains and losses, respectively. The CPT valuation of the prospect is given by:

$$w^-(p)v(x) + w^+(q)v(y)$$

for mixed prospects, $x < 0 < y$

$$[w^+(p+q) - w^+(q)]v(x) + w^+(q)v(y)$$

for pure gain prospects, $0 \le x < y$

$$[w^-(p+q) - w^-(q)]v(x) + w^-(q)v(y)$$

for pure loss prospects, $y < x \le 0$.

The first equation illustrates sign dependence: a different weighting function is applied separately to the loss and gain portions of mixed prospects. The second and third equations illustrate rank dependence for gains and losses, respectively: extreme (y) outcomes are weighted by the impact of their respective probabilities, whereas intermediate outcomes (x) are weighted by the difference in impact of the probability of receiving an outcome at least as good as x and the impact of the probability of receiving an outcome that is strictly better than x. A more general characterization of CPT that applies to any finite number of outcomes and decisions under uncertainty is included in the Appendix to this chapter.

For decision under risk, the predictions of CPT coincide with OPT for all two-outcome risky prospects and all mixed (gain–loss) three-outcome prospects[4] when one outcome is zero, assuming $w^+ = w^-$. Because elicitation of prospect theory parameters (reviewed in the following section) usually requires the use of two-outcome prospects, we illustrate how they coincide for a two-outcome (pure gain) prospect below. Consider a prospect $(x, p; y)$ that offers \$$x$ with probability p and otherwise \$$y$, where $x > y$. According to CPT:

$$V(x, p; y) = [1 - w(p)]v(y) + w(p)\,v(x).$$

According to OPT, decision makers tend to invoke the editing operation of *segregation*, treating the smaller outcome y as a certainty, and reframing the prospect as a p chance of getting an additional $x - y$. Thus, we get:

$$V(x, p; y) = v(y) + w(p)[v(x) - v(y)]$$

which can be rearranged into the same expression as above. It is also easy to see that when $y = 0$, $V(x, p) = w(p)\,v(x)$ under both CPT and OPT.

Decision Weights Under Risk Versus Uncertainty: the Two-stage Model

As we have seen, the risky weighting function is assumed to exhibit greater sensitivity to changes in probability (i.e. higher slope) near the natural boundaries of 0 and 1 than in the midpoint of the scale. A characterization of the weighting function that generalizes

---

[2] Wu and Markle (2008) document systematic violations of gain–loss separability. Their results suggest different weighting function parameter values for mixed (gain–loss) prospects than for single domain (pure gain or pure loss) prospects.

[3] Rank-dependence is motivated in part by the concern that non-linear decision weights applied directly to multiple simple outcomes can give rise to violations of stochastic dominance. For instance, a prospect that offers a .01 chance of \$99 and a .01 chance of \$100 might be preferred to a prospect that offers a .02 chance of \$100 due to the overweighting of low probabilities, even though the latter prospect dominates the former prospect. OPT circumvents this problem for simple prospects by assuming that transparent violations of dominance are eliminated in the editing phase; CPT handles this problem through a rank-dependent decision weights that sum to one for pure gain or loss prospects. For further discussion of advantages of CPT over OPT when modeling preferences involving complex prospects, see Fennema and Wakker, 1997.

[4] Gonzalez and Wu (2003) estimated prospect theory weighting functions and value functions obtained from cash equivalents for two-outcome gambles, in which OPT and CPT coincide, and applied these estimates to predict cash equivalents for three-outcome gambles, in which they do not. Interestingly, they found systematic over-prediction for OPT and systematic under-prediction for CPT.

this observation from risk to uncertainty through the measure of "bounded subadditivity" is presented in Tversky and Fox (1995; see also Tversky and Wakker, 1995; Wu and Gonzalez, 1999). Informally, bounded subadditivity quantifies a decision-maker's diminished sensitivity to events when they are added to or subtracted from intermediate events compared to when they are added to impossibility or subtracted from certainty.

Several studies suggest that decisions under uncertainty accord well with a two-stage model in which participants first judge likelihood of events on which outcomes depend, then apply the inverse S-shaped weighting function to these probabilities, consistent with prospect theory (Tversky and Fox, 1995; Fox and Tversky, 1998; for a theoretical treatment, see Wakker, 2004). That is, the uncertain decision weight W of event E is given by

$$W(E) = w(P(E))$$

where $P(E)$ is the (non-additive) judged probability of event E and $w(\cdot)$ is the risky weighting function. For instance, consider the prospect "win \$100 if the Lakers beat the Celtics." A person's decision weight of "Lakers beat the Celtics" can be predicted well from his risky weighting function applied to his judged probability of the event "Lakers beat the Celtics." Judged probabilities are assumed to accord with support theory (Tversky and Koehler, 1994; Rottenstreich and Tversky, 1997), a behavioral model that conceives of judged probability as the proportion of support that a person associates with a focal hypothesis (e.g., that the Lakers will win) against its complement (the Celtics will win). Fox and Tversky (1998) review several studies that demonstrate the predictive validity of the two-stage model (see also Wu and Gonzalez, 1999; Fox and See, 2003; but see too Kilka and Weber, 2001).

Ambiguity Aversion and Source Preferences

Decisions under uncertainty can be further complicated by preferences to bet on particular sources of uncertainty. Ellsberg (1961) observed that people prefer to bet on events with known rather than unknown probabilities, a phenomenon known as ambiguity aversion (for a review, see Camerer and Weber, 1992; see also Fox and See, 2003). This phenomenon may partially explain, for example, the common preference to invest in the domestic stock market and under-diversify into foreign markets (French and Poterba, 1991). Ambiguity aversion appears to be driven by reluctance to act in situations in which a person feels comparatively ignorant of predicting outcomes (Heath and Tversky, 1991), and such preferences tend to diminish or disappear in the absence of a direct comparison between more and less familiar events or with more or less knowledgeable individuals (Fox and Tversky, 1995; Chow and Sarin, 2001; Fox and Weber, 2002). For a discussion of how source preferences can be incorporated into the two-stage model, see Fox and Tversky (1998).

Decisions from Experience

Finally, situations in which people learn relative frequencies of possible outcomes from experience (e.g., as in the Iowa Gambling Task or Balloon Analog Risk Task), learning can be complicated by sampling error. In particular, according to the binomial distribution very rare events are generally more likely to be under-sampled than over-sampled, and the opposite is true for very common events. For instance, imagine a situation in which a decision maker samples outcomes from two decks of cards: the first deck offers a .05 chance of \$100 (and nothing otherwise) while the second deck offers \$5 for sure. If decision makers sample a dozen cards from each deck, most will *never* sample \$100 from the first deck and therefore face an apparent choice between \$0 for sure and \$5 for sure, and therefore forego the 5% chance of \$100, contrary to the pattern observed in decision under risk. (For further discussion of these issues, see Hertwig *et al.*, 2004; Fox and Hadar, 2006). For further discussion of how the two-stage model can be extended to situations in which outcomes are learned from experience, see Hadar and Fox (2008).

PROSPECT THEORY MEASUREMENT

Several applications of prospect theory – from neuroeconomics to decision analysis to behavioral finance – require individual assessment of value and weighting functions. In order to measure the shape of the value and weighting functions exhibited by participants in the laboratory, we must first discuss how these functions can be formally modeled. We next discuss procedures for eliciting values and decision weights.

Parameterization

It is important to note that, in prospect theory, value and weighting functions are characterized by their qualitative properties rather than particular functional forms. It is often convenient, however, to fit data to equations that satisfy these qualitative properties. A survey of parameterizations of prospect theory's value and weighting functions can be found in Stott (2006). We review below the functional forms that have received the most attention in the literature to date.

Value Function

The value function is assumed to be concave for gains, convex for losses, and steeper for losses than for gains. By far the most popular parameterization, advanced by Kahneman and Tversky (1992) relies on a power function:

$$v(x) = \begin{cases} x^{\alpha} & x \geq 0 \\ -\lambda(-x)^{\beta} & x < 0 \end{cases} \qquad \text{(V1)}$$

where α, $\beta > 0$ measure the curvature of the value function for gains and losses, respectively, and λ is the coefficient of loss aversion. Thus, the value function for gains (losses) is increasingly concave (convex) for smaller values of $\alpha(\beta) < 1$, and loss aversion is more pronounced for larger values of $\lambda > 1$. Tversky and Kahneman (1992) estimated median values of $\alpha = .88$, $\beta = .88$, and $\lambda = 2.25$ among their sample of college students. In prospect theory the power function is equivalent to *preference homotheticity*: as the stakes of a prospect (x, p) are multiplied by a constant k, then so is the certainty equivalent of that prospect, $C(x, p)$ so that $C(kx, p) = kC\ (x, p)$. (see, e.g., Tversky, 1967). Empirically this assumption tends to hold up only within an order of magnitude or so, and as the stakes of gambles increase by orders of magnitude, risk aversion tends to increase for gains – especially when the stakes are real (Holt and Laury, 2002); the evidence for losses is mixed (Fehr-Duda *et al.*, 2007). Thus, for example, a person who is indifferent between \$3 and (\$10, .5) will tend strictly to prefer \$30 over (\$100, .5). Nevertheless, most applications of prospect theory have assumed a power value function. Other common functional forms include the logarithmic function $v(x) = \ln\ (\alpha + x)$, originally proposed by Bernoulli (1738), which captures the notion that marginal utility is proportional to wealth, and quadratic $v(x) = \alpha x - x^2$, which can be reformulated in terms of a prospect's mean and variance, which is convenient in finance models. (For a discussion of additional forms including exponential and expo-power, see Abdellaoui *et al.*, 2007a.)

Surprisingly, there is no canonical definition or associated measure of loss aversion, though several have been proposed. First, in the original formulation of prospect theory (Kahneman and Tversky, 1979), loss aversion was defined as the tendency for the negative value of losses to be larger than the value of corresponding gains (i.e., $-v(-x) > v(x)$ for all $x > 0$) so that a coefficient of loss aversion might be defined, for example, by the mean or median value of $-v(-x)/v(x)$ over a particular range of x. Second, the aforementioned parameterization (V1) from Tversky and Kahneman (1992) that assumes a power value function implicitly defines the loss aversion as the ratio of value of losing a dollar to gaining a dollar (i.e., $-v(-\$1) > v(\$1)$) so that the coefficient is defined by $-v(-\$1)/v(\$1)$. Third, Wakker and Tversky (1993) defined loss aversion as the requirement that the slope of the value function for any amount lost is larger than the slope of the value function for the corresponding amount gained (i.e., $v'(-x) > v'(x)$) so that the coefficient can be defined by the mean or median value of $v'(-x)/v'(x)$. Note that if one assumes a simplified value function that is piecewise linear (as in, for example, Tom *et al.*, 2007), then all three of these definitions coincide. For a fuller discussion, see Abdellaoui *et al.* (2007b).

Weighting Function

In fitting their data, Tversky and Kahneman (1992) asserted a single-parameter weighting function:

$$w(p) = p^{\gamma}/(p^{\gamma} + (1-p)^{\gamma})^{1/\gamma}. \qquad \text{(W1)}$$

This form is inverse-S shaped, with overweighting of low probabilities and underweighting of moderate to high probabilities for values of $\gamma < 1$. This function is plotted for various values of γ in Figure 11.3A.

Perhaps the most popular form of the weighting function, due to Lattimore *et al.* (1992; see also Goldstein and Einhorn, 1987) assumes that the relation between w and p is linear in a log-odds metric:

$$\ln\frac{w(p)}{1-w(p)} = \gamma\ln\frac{p}{1-p} + \ln\delta$$

which reduces to

$$w(p) = \frac{\delta p^{\gamma}}{\delta p^{\gamma} + (1-p)^{\gamma}} \qquad \text{(W2)}$$

where $\delta > 0$ measures the elevation of the weighing function and $\gamma > 0$ measures its degree of curvature. The weighting function is more elevated (exhibiting less overall risk aversion for gains, more overall risk aversion for losses) as δ increases and more curved (exhibiting more rapidly diminishing sensitivity to probabilities around the boundaries of 0 and 1) as $\gamma < 1$ decreases (the function exhibits an S-shaped pattern that is more pronounced for larger values of $\gamma > 1$). Typically, the decision weights of complementary events sum to less than one ($w(p) + w(1-p) < 1$), a property known as *subcertainty* (Kahneman and Tversky, 1979). This property is satisfied whenever $\delta < 1$. The Lattimore function is plotted for various values of the elevation parameter δ and curvature parameter γ in Figures 11.3b and 11.3c, respectively.

Prelec (1998; see also 2000) derived a functional form of the weighting function that accommodates three principles: (1) overweighting of low probabilities and underweighting of high probabilities; (2) subproportionality of decision weights (a condition that derives from the common ratio effect, decisions 1 and 2 above); and (3) sub-additivity of decision weights (a condition that derives from the common consequence effect, decisions 3 and 4 above). These three principles are all subsumed by a single axiom called *compound invariance*[5] which implies the following functional form of the weighting function:

$$w(p) = \exp[-\delta(-\ln p)^{\gamma}] \qquad \text{(W3A)}$$

where $\delta, \gamma > 0$. When $\delta = 1$, Prelec's function collapses to a single-parameter form:

$$w(p) = \exp[-(-\ln p)^{\gamma}] \qquad \text{(W3B)}$$

which implies a weighting function that crosses the identity at $1/e$. Prelec's two-parameter function is plotted for various values of the elevation parameter δ in Figure 11.3d, and the one-parameter function (i.e., $\delta = 1$) is plotted for various values of the curvature parameter γ in Figure 11.3e.

The prospect theory value and weighting function parameters can all be estimated for individuals using simple choice tasks on computer. Table 11.3 presents measured parameters for monetary gambles from several studies that have assumed a power value function and various weighting functions described above.

Although the typical measured values of these parameters suggest an S-shaped value function ($0 < \alpha, \beta < 1$) with loss aversion ($\lambda > 1$), and an inverse-S shaped weighting function that crosses the identity line below .5, there is considerable heterogeneity between individuals in these measured parameters. For instance, in a sample of 10 psychology graduate students evaluating gambles involving only the possibility of gains, Gonzalez and Wu (1999) obtained measures of α in the range from .23 to.68 (V1), δ in the range from .21 to 1.51, and γ in the range from .15 to .89 (W2).

As a practical matter, although the two-parameter functions (W2) and (W3) have different axiomatic implications, they are difficult to distinguish empirically in the normal range (i.e., .01 to .99) of probabilities (see Gonzalez and Wu, 1999). For the remainder of the chapter, we will refer to the parameters from the Lattimore *et al*. (1992) function (W2).

Interaction of $v(\cdot)$ and $w(\cdot)$

As mentioned above, prospect theory value and weighting functions both contribute to observed risk attitudes: concavity (convexity) of the value function contributes to risk aversion (seeking) for pure gain (loss) prospects that is reinforced by underweighting of moderate to high probabilities and reversed by overweighting of low probabilities; loss aversion contributes to risk aversion for mixed prospects. To see more clearly how the value and weighting functions interact, consider the simple case of a prospect (x, p) that offers \$$x$ with probability p (and nothing otherwise). Let $c(x, p)$ be the certainty equivalent of (x, p). For instance, a decision maker for whom $c(100, .5) = 30$ is indifferent between receiving \$30 for sure or 50–50 chance of \$100 or nothing. Thus, this decision maker would strictly prefer the prospect to \$29 and would strictly prefer \$31 to the prospect. If we elicit certainty equivalents for a number of prospects in which we hold x constant and vary p, then we can derive a plot of *normalized certainty equivalents*, c/x as a function of probability. Such a plot can be instructive, because it indicates probabilities (of two-outcome gambles) for which the decision maker is risk seeking ($c/x > p$), risk neutral ($c/x = p$), and risk averse ($c/x < p$) by whether the curve lies above, on, or below the identity line, respectively.

To see how $w(\cdot)$ and $v(\cdot)$ jointly contribute to risk attitudes, note that, under prospect theory, $V(c) = V(x, p)$, so that $v(c) = w(p)v(x)$ or $w(p) = v(c)/v(x)$. Assuming the power value function (V1), we get $w(p) = (c/x)^{\alpha}$, or

$$c/x = w(p)^{1/\alpha}.$$

In the case of gains, normalized certainty equivalents will increase with the parameter α and, assuming a concave value function ($\alpha < 1$) that is correctly measured, they will be lower than corresponding decision weights. These observations give rise to two important implications. First, overweighting of low probabilities does not necessarily translate into risk-seeking for low-probability gains. To illustrate, consider the weighting function obtained from the median data of Gonzalez and

---

[5] Defined as: for any outcomes x, y, x', y', probabilities q, p, r, s, and the compounding integer $N \geq 1$, if $(x, p) \sim (y, q)$ and $(x, r) \sim (y, s)$ then $(x', p^N) \sim (y', q^N)$ implies $(x', r^N) \sim (y', s^N)$.

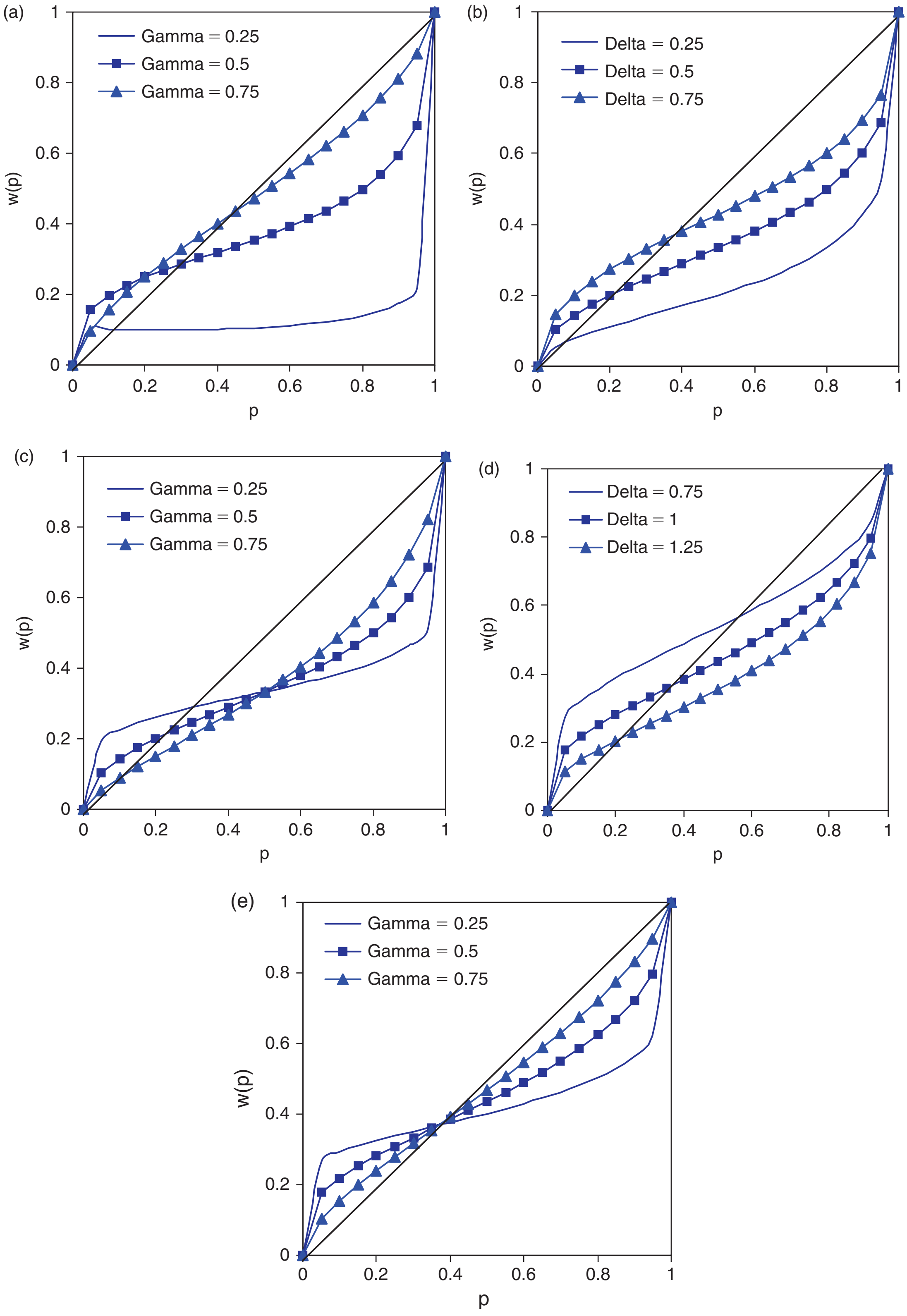


FIGURE 11.3 Most common parametric forms used for modeling the probability weighting function from prospect theory. (a) Tversky and Kahneman's (1992) function for various values of γ (W1). (b) Lattimore *et al.*'s (1992) function for various values of δ assuming $\gamma = .5$ (W2). (c) Lattimore *et al.*'s (1992) function for various values of γ assuming $\delta = .5$ (W2). (d) Prelec's (1998) function for various values of δ assuming $\gamma = .5$ (W3A). (e) Prelec's (1998) function for various values of γ assuming $\delta = 1$ (W3B).

TABLE 11.3 Measured value function parameters for money from several studies

Functional form	Study	Subject population	Parameter estimates
(a)			
(V1) $v(x) = \begin{cases} x^\alpha & x \geq 0 \\ -\lambda(-x)^\beta & x < 0 \end{cases}$	Tversky and Kahneman (1992)	n = 25 graduate students (median fitted parameters)	$\alpha = .88$ $\beta = .88$ $\lambda = 2.25$
	Camerer and Ho (1994)	Weighted average of nine studies reviewed	$\alpha = .23$
	Wu and Gonzalez (1996)	n = 420 undergraduates (fitted to binary choice data)	$\alpha = .49$
	Gonzalez and Wu (1999)	n = 10 psychology graduate students (median data)	$\alpha = .49$
	Abdellaoui (2000)	n = 46 economics students (median data)	$\alpha = .89$ $\beta = .92$
	Etchart-Vincent (2004)	n = 35 business students (median data)	$\beta = .97$
	Abdellaoui *et al.* (2005)	n = 41 business graduate students (median fitted parameters)	$\alpha = .91$ $\beta = .96$
	Stott (2006)	n = 96 university students (median fitted data)	$\alpha = .19$
	Abdellaoui *et al.* (2007b)	n = 48 economics students (median data)	$\alpha = .75$ $\beta = .74$
	Abdellaoui *et al.* (2007c)	n = 48 economics and math graduate students (median data)	$\alpha = .86$ $\beta = 1.06$ $\lambda = 2.61$
(b)			
(W1) $W(p) = p^\gamma/(p^\gamma + (1-p)^\gamma)^{1/\gamma}$	Tversky and Kahneman (1992)	n = 25 graduate students (median fitted parameters)	$\gamma^+ = .61$ $\gamma^- = .69$
	Camerer and Ho (1994)	Weighted average of nine studies reviewed	$\gamma^+ = .56$
	Wu and Gonzalez (1996)	n = 420 undergraduates (fitted to binary choice data)	$\gamma^+ = .71$
	Abdellaoui (2000)	n = 46 economics students (median data)	$\gamma^+ = .60$ $\gamma^- = .70$
	Stott (2006)	n = 96 university students (median fitted data)	$\gamma^+ = .96$
(W2) $w(p) = \dfrac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma}$	Tversky and Fox (1995)	n = 40 student football fans (median data, with $\alpha = .88$)	$\gamma^+ = .69$ $\delta^+ = .77$
	Wu and Gonzalez (1996)	n = 420 undergraduates (fitted to binary choice data)	$\gamma^+ = .68$ $\delta^+ = .84$
	Gonzalez and Wu (1999)	n = 10 psychology graduate students (median data)	$\gamma^+ = .44$ $\delta^+ = .77$
	Abdellaoui (2000)	n = 46 economics students (median data)	$\gamma^+ = .60$ $\delta^+ = .65$ $\gamma^- = .65$ $\delta^- = .84$
	Abdellaoui *et al.* (2005)	n = 41 business graduate students (median data)	$\gamma^+ = .83$ $\delta^+ = .98$ $\gamma^- = .84$ $\delta^- = 1.3$
	Stott (2006)	n = 96 university students (median fitted data)	$\gamma^+ = 1.4$ $\delta^+ = .96$
(W3A) $w(p) = \exp[-\delta(-\ln p)^\gamma]$	Stott (2006)	n = 96 university students (median fitted data)	$\gamma^+ = 1.0$ $\delta^+ = 1.0$
(W3B) $w(p) = \exp[-(-\ln p)^\gamma]$	Wu and Gonzalez (1996)	n = 420 undergraduates (fitted to binary choice data)	$\gamma^+ = .74$
	Stott (2006)	n = 96 university students (median fitted data)	$\gamma^+ = 1.0$

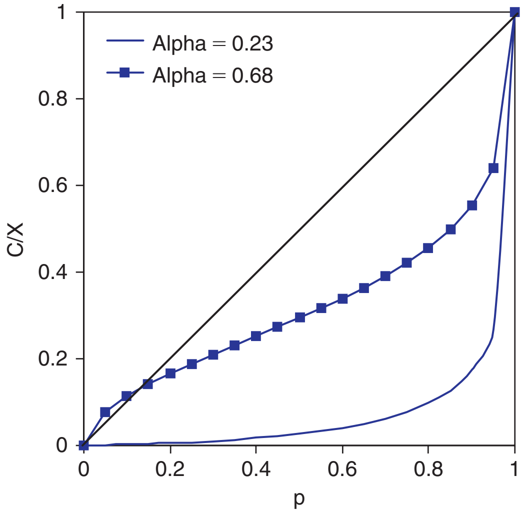


FIGURE 11.4 Normalized certainty equivalents as a function of probability assuming the Lattimore weighting function, with $\delta = .77$ and $\gamma = .44$ (median values from Gonzalez and Wu, 1999) and assuming a power value function, with $\alpha = .23$ and .68 (the range obtained from participants of Gonzalez and Wu, 1999). This figure illustrates the interaction of the value and weighting functions in determining risk attitudes.

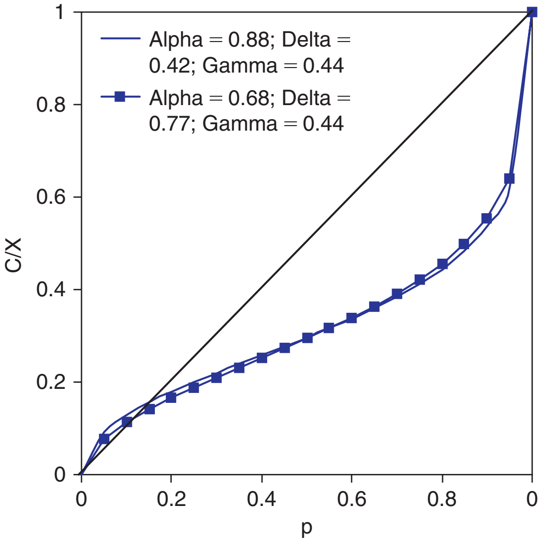


FIGURE 11.5 Normalized certainty equivalents as a function of probability assuming the Lattimore weighting function and power value function with $\alpha = .68$, $\delta = .77$, and $\gamma = .44$. versus $\alpha = .88$, $\delta = .42$, and $\gamma = .44$. This figure illustrates the difficulty empirically distinguishing between elevation of the weighting function and curvature of the value function.

Wu (1999), assuming the Lattimore *et al.* (1992) function (W2), with $\delta = .77$, $\gamma = .44$, which illustrates considerable overweighting of low probabilities; for example, $w(.05) = .17$. In that study, the authors obtained α in the range from .68 (moderate concavity) to .23 (extreme concavity) for their ten participants. Using these extreme values, we obtain wildly different c/x functions as depicted in Figure 11.4. For instance, given these values $c(100, .05) = 7.65$ and .05, respectively, indicating moderate risk-seeking and extreme risk aversion, respectively.

Second, the interaction of value- and weighting-functions makes it difficult empirically to distinguish variations in the measured elevation of the weighting function from variations in the measured curvature of the value function. For instance, as mentioned above, when $\alpha = .68$, $\delta = .77$, and $\gamma = .44$. we get $c(100, .05) = 7.65$. This same certainty equivalent follows assuming, for example, $\alpha = .88$, $\delta = .42$, and $\gamma = .44$. Both of these normalized certainty equivalent functions are illustrated in Figure 11.5. Thus, if one is concerned with parsing the contribution of subjective value versus probability weighting on observed risk attitudes, it is important to elicit the value and weighting functions with care. For instance, if one assumes a single parameter weighting function (e.g., (W1) or (W3B)) when "true" weighting functions vary in their elevation, incorrect measures may be obtained. A researcher may believe that a particular pattern of neural activity covaries with curvature of the value function, when in fact it covaries with elevation of the weighting function.

Elicitation

Several methods have been proposed for eliciting value and weighting function parameters. Broadly speaking, these methods fall into four categories:

1. Statistical methods that estimate $v(x_i)$ and $w(p_i)$ from a participant's cash equivalents for prospects that factorial combine each x_i, and p_i.
2. Non-parametric methods that separately assess values then assess decision weights, making no assumptions concerning the functional form of the value-and weighting-functions.
3. Semi-parametric methods that assume a functional form for the value- or weighting-function and assess the other function non-parametrically.
4. Parametric methods that assume a functional form of both the value and weighting functions.

We will review each of these methods in turn then evaluate their relative strengths and weaknesses.

Statistical Method: Gonzalez and Wu (1999)

Perhaps the most careful elicitation method of prospect theory value and weighting functions to date was advanced by Gonzalez and Wu (1999). Ten graduate students in Psychology from the University of Washington were paid \$50 plus an incentive-compatible payment (contingent on their choices) for their participation in four 1-hour sessions on computer[6]. Participants were presented with 15 two-outcome (non-negative) gambles crossed with 11 probabilities (=165 gambles), presented in a random order.

Certainty equivalents were assessed for each gamble through a series of choices. For instance, consider the prospect that offered a 50–50 chance of \$100 or nothing. A participant was asked if he preferred to receive the prospect or various sure amounts that ranged from \$100 to \$0 in increments of \$20. If a participant indicated that he preferred \$40 for sure over the prospect but preferred the prospect over \$20 for sure, then a second round of choices would be presented that spanned this narrower range (from \$40 to \$20). This process was repeated until certainty equivalents could be estimated to the nearest dollar. If, for example, a participant indicated a preference for a sure \$36 over the prospect but a preference for the prospect over a sure \$35, then the researchers estimated $c(100, .5) = 35.5$.

The estimation procedure used by Gonzalez and Wu (1999) was non-parametric in that it did not make any assumptions concerning the functional form of the $v(\cdot)$ or $w(\cdot)$. Their algorithm treated the value of each of the possible outcomes and the weight of each of the probabilities presented as a parameter to be estimated. These parameters were estimated using an alternating least squares procedure in which each step either held w constant and estimated v or held v constant and estimated w. The authors assert that this analysis converged on parameter estimates relatively quickly.

The statistical method of Gonzalez and Wu (1999) has several advantages over alternative methods. The elicitation is not very cognitively demanding, as participants are merely required to price two-outcome gambles. The procedure gives rise to estimates of values and decision weights that are not distorted by parametric misspecification. On the other hand, the procedure is demanding of participants' time as it requires pricing of a large number of gambles to get stable estimates (the original study required participants to assess 165 two-outcome gambles, each through a series of several choices). The procedure has not yet been applied to the domain of losses or mixed prospects, but such an extension would be straightforward.

Non-parametric Methods

Several other fully non-parametric methods have been advanced for analytically assessing $v(\cdot)$ and $w(\cdot)$. All of them rely on a two-stage process in which $v(\cdot)$ is assessed in a first phase, then applied to the measurement of $w(\cdot)$. The most popular approach to assessing values that makes no assumptions concerning the weighting of probabilities is the *tradeoff method* (Wakker and Deneffe, 1996). The tradeoff method requires participants to make choices between two two-outcome prospects $(x, p; y)$ that offer \$$x$ with probability p otherwise \$$y$, with one of the outcomes adjusted following each choice until indifference between the gambles can be established. Consider a pair of reference outcomes $R > r$, a pair of variable outcomes $x_1 > x_0$, and a fixed probability p. On each trial the values of R, r, x_0, and p are fixed, and x_1 is varied until the participant reveals that

$$(x_1, p; r) \sim (x_0, p; R).$$

For instance, a participant might be offered a choice between a 50–50 chance of \$100 or \$20 versus a 50–50 chance of \$70 or \$40. If the participant prefers the latter gamble, then the variable payoff of the first gamble (\$100) adjusts to a higher amount (say, \$110). The variable amount can be raised or lowered by decreasing increments until the participant confirms that both prospects are equally attractive. Once indifference is established for this first pair of prospects, the procedure is repeated for a second pair of prospects with the same probability and reference outcomes, but a new variable outcome $x_2 > x_1$, until it is established that:

$$(x_2, p; r) \sim (x_1, p; R).$$

According to CPT[7], the first indifference gives us

$$v(r)[1 - w(p)] + v(x_1)w(p) = v(R)[1 - w(p)] \\ + v(x_0)w(p)$$

---

[6] An incentive-compatible payoff is a payment contingent on choice that encourages honest responses by participants. Experimental economists are generally skeptical of results of studies that do not include such incentives whereas experimental psychologists generally put more credence into responses to purely hypothetical choices. In practice, the addition of incentives tends to reduce noise in participant responses and may lead to decreased framing effects and greater risk aversion (for reviews, see Camerer and Hogarth, 1999; Hertwig and Ortmann, 2001).

[7] Assuming $x_0 > R$; this result can be relaxed without affecting the result of the elicitation.

so that

$$w(p)[v(x_1) - v(x_0)] = [1 - w(p)][v(R) - v(r)]$$

and the second indifference gives us

$$v(r)[1 - w(p)] + v(x_2)w(p) = v(R)[1 - w(p)] + v(x_1)w(p)$$

so that

$$w(p)[v(x_2) - v(x_1)] = [1 - w(p)][v(R) - v(r)].$$

Together these indifferences imply equal value intervals as follows:

$$v(x_1) - v(x_0) = v(x_2) - v(x_1).$$

Setting $x_0 = 0$ and $v(x_0) = 0$, we get $v(x_2) = 2v(x_1)$. By eliciting similar yoked indifferences to obtain x_3, x_4, etc., we can generate a standard sequence of outcomes that are spaced equally in subjective value space to construct a parameter-free value function for gains. A similar exercise can be repeated in the measurement of the value function for losses (for an example in the domain of losses, see Fennema and van Assen, 1999).

Once a measure of several values has been obtained from a participant, one can proceed to measure decision weights non-parametrically. Arguably the most popular method, advanced by Abdellaoui (2000), uses the standard sequence of outcomes x_0, ..., x_n to elicit a standard series of probabilities $p_1, \ldots, p_{n-1}$ that are equally spaced in terms of their decision weights. This is done by eliciting probabilities such that a mixture of the highest and lowest outcome in the standard sequence is equally attractive to each of the internal outcomes in that sequence. Thus, by establishing for each x_i ($i = 1, \ldots, n - 1$) the following indifference:

$$(x_n, p_i; x_0) \sim x_i.$$

CPT implies:

$$w(p_i) = \frac{v(x_i) - v(x_0)}{v(x_n) - v(x_0)}.$$

Because the values of x_i were constructed, using the tradeoff method, to be equally spaced in terms of their expected value, the above equation reduces to:

$$w(p_i) = i/n.$$

An analogous procedure can be followed for losses.

Bleichrodt and Pinto (2000) advanced a similar two-step procedure that first relies on the tradeoff method to elicit a standard sequence of outcomes, then elicits decision weights through a matching procedure. Instead of eliciting probabilities that lead to indifference between prospects, their method fixes probabilities and elicits outcomes that match pairs of two-outcome prospects[8]. Such a procedure was used to measure the weighting function for losses by Etchart-Vincent (2004). Another similar method has recently been proposed by van de Kuilen *et al.* (2006), though in an experiment this method yielded a weighting function for gains that was convex rather than the customary inverse-S shape (concave then convex).

The aforementioned non-parametric elicitations can be used to assess value- and weighting-functions separately for gains and losses. Because the value function is a ratio scale (unique to multiplication by a positive constant) a separate procedure using mixed (gain−loss) gambles is required to assess loss aversion. A parameter-free procedure has been advanced by Abdellaoui *et al.* (2007b). Details of the procedure are beyond the scope of this chapter, but the gist is as follows. The first step entails determining, through a series of indifferences between prospects, the probabilities p_g and p_l for which $w^+(p_g)$ and $w^-(p_l) = 1/2$. This allows determination, in a second stage, of outcome amounts that are midpoints in value space for losses. The third stage links value for losses and gains through a series of indifferences that determines a gain outcome that is the mirror image of a loss outcome in value space (i.e., has the same absolute value of utility/value). Finally, the fourth step repeats the second step by determining outcomes that are midpoints in value space for gains. The method of Abdellaoui *et al.* (2007b) is mathematically elegant and yielded clean results consistent with prospect theory in the analysis of aggregate data from a sample of 48 economics students. However, the task is cognitively demanding, as it involves choices between pairs of two-outcome gambles, and laborious, as it entails a complex four-step procedure.

Non-parametric methods tend to be less time consuming than statistical methods of elicitation. Also, unlike semi-parametric and fully parametric methods, they make no assumptions concerning the functional form of the value and weighting functions that might distort measurement, though functions can be fitted to the measured values and weights that are obtained.

---

[8] Note that because the new outcomes may not be included in the standard sequence this method requires an interpolation procedure and thus is not fully non-parametric.

Moreover, non-parametric methods preserve a direct link between specific choices and measured utilities so that specific inconsistencies can be traced to particular choices. Unfortunately, non-parametric methods are generally quite cognitively demanding, requiring choices between multiple two-outcome prospects (or even more complicated choices). Thus, these methods may not give utterly robust measurements, as participants may fall back on decision heuristics (such as expected value maximization) or respond in an inconsistent manner. Moreover, because these methods generally rely on elicitation of a standard sequence of values using the tradeoff method, there is the possibility that error in measuring the first step in the sequence will be propagated throughout the measurement of values and therefore lead to further error in the measurement of decision weights (however, studies that have investigated error propagation have thus far found no large effect; see Bleichrodt and Pinto, 2000; Abdellaoui *et al.*, 2005). Note that only methods listed as allowing simultaneous measurement of both v^+ and v^- can also allow measurement of loss aversion.

Semi-Parametric Methods

Semi-parametric elicitation methods assume a parametric form of the value function in order to derive non-parametric estimates of decision weights. The simplest semi-parametric approach is to assume a power value function, $v(x) = x^\alpha$, as fitted to non-parametric measurement of value using the tradeoff method (or assuming representative parameters from previous studies of similar participant populations). Next, decision weights for various probabilities can be determined by eliciting certainty equivalents $c(x, p_i)$ for prospects that pay a fixed amount x with probabilities p_i. According to prospect theory, $c(x, p_i)^\alpha = w(p_i)x^\alpha$. Thus, each decision weight is given by:

$$w(p_i) = [c(x, p_i)/x]^\alpha.$$

Of course, this method depends on the accuracy of the first-stage measurement of utility.

A more elegant semi-parametric method was recently advanced by Abdellaoui *et al.* (2007c). This method entails three stages. In the first stage, the value function for gains is elicited and decision weights are measured parameters. This is done by eliciting certainty equivalents G_i for a series of prospects $(x_i, p_g; y_i)$ $(x_i > y_i \geq 0, i = 1, ..., k)$. According to CPT:

$$v(G_i) = v(y_i)[1 - w(p_g)] + v(x_i)w(p_g).$$

Define $w(p_g) \equiv \omega^+$ and assume a power value function $v(x) = x^\alpha$. We get:

$$G_i = (w^+(x_i^\alpha - y_i^\alpha) + y_i^\alpha)^{1/\alpha}.$$

Thus, by varying x_i and y_i and measuring cash equivalents G_i, the parameters ω^+ and α can be estimated using non-linear regression. An analogous method can be used for the domain of losses to measure ω^-, the decision weight of losing with probability $p_l = 1 - p_g$, and β, the power value coefficient for losses. Finally, a third stage links the value function for gains and losses by selecting a gain amount G^* within the range of value measured in step 1, then determining the loss amount L^* such that a participant finds the mixed prospect $(G^*, p_g; L^*)$ barely acceptable (i.e., is indifferent to playing the prospect or not). This implies that:

$$w^+v(G^*) + w^-\lambda v(L^*) = v(0) = 0$$

so one can easily solve for λ. Although the method of Abdellaoui *et al.* (2007c) is designed to elicit value function and loss aversion parameters, it also provides as a byproduct measurement of a decision weight. By repeating the procedure for various probabilities of gain and loss, several decision weights can be obtained for mapping more complete weighting functions.

Semi-parametric methods provide a compromise between accuracy of a non-parametric elicitation method and the efficiency of a parametric method. They tend to be less cognitively demanding and less time consuming than pure non-parametric methods and the statistical method.

Parametric Methods

The final method for eliciting prospect theory value- and weighting-functions is a purely parametric approach. Tversky and Kahneman (1992) elicited cash equivalents for a number of single- and two-outcome prospects entailing pure gains, pure losses, and mixed outcomes. These were entered into a non-linear regression assuming a power value function (V1) and single-parameter weighting function (W1).

A simpler procedure can be executed using Prelec's (1998) single-parameter weighting function (W3B) and a power value function. If we elicit a number of certainty equivalents c_{ij} for prospects that pay \$$x_i$ with probability p_j, then we get by prospect theory:

$$c_{ij}{}^\alpha = x_i{}^\alpha \exp[-(\ln p)^\gamma].$$

TABLE 11.4 Major elicitation methods

Method class	Reference	Prospect theory component(s)	Cognitive demands	Time required
Statistical	Gonzalez and Wu (1999)	All	Low	High
Non-parametric	Wakker and Deneffe (1996)	v^+ or v^-	High	Medium
	Abdellaoui *et al*. (2007b)	v^+ and v^-	High	Medium
	Abdellaoui (2000)	w^+ or w^-	High	Medium
	Bleichrodt and Pinto (2000)	w^+ or w^-	High	Medium
Semi-parametric	Abdellaoui *et al*. (2007c)	v^+ and v^-, limited w^+, w^-	Medium	Low
Parametric	Prelec (1998)	v^+, w^+ or v^-, w^-	Low	Medium

Collecting outcomes on the left side of the equation and taking the double log of both sides, we get:

$$-\ln[-\ln(c_{ij}/x_i)] = \ln(\alpha) + \gamma[-\ln(-\ln p_j)].$$

This equation lends itself to linear regression to determine the parameters α and γ.

Parametric estimation of value and weighting functions has several advantages over other methods. The task of pricing simple prospects is cognitively tractable, the time requirements are relatively small, and this method tends to yield relatively reliable measurement. On the other hand, this method is susceptible to parametric misspecification, particularly if one assumes a single parameter weighting function (as in the method of Prelec described above) so that it is difficult to distinguish the curvature of the value function from elevation of the weighting function.

Table 11.4 summarizes the major methods for prospect theory elicitation, listing strengths and weaknesses of each method. All entail tradeoffs, and the particular method used by researchers will be determined by the cognitive sophistication of participants, time constraints, and technical constraints of the study in question.

Determining Certainty Equivalents

Several elicitation methods discussed above require determination of certainty equivalents of various prospects. The most straightforward (but cognitively demanding) method is to elicit them directly by asking participants for the sure amount of money c that they find equally attractive to a prospect (x, p). Participants can be provided incentives for accuracy using the method described by Becker *et al.* (1964)[9]. Alternatively, one might ask participants for the probability p such that they find the prospect (x, p) equally attractive to the sure amount c. Empirically such elicitations tend to be quite noisy, but they are quick and convenient.

We caution researchers against such direct matching procedures. Prospect theory was originally articulated as a model of simple choice between prospects. Direct elicitation of sure amounts or probabilities to match prospects relies on the assumption of *procedure invariance*: two strategically equivalent methods of assessing preference should lead to the identical orderings between prospects. Unfortunately, this assumption is routinely violated. First, people generally afford more weight to outcomes relative to probabilities when they price prospects than when they choose between them. This can give rise to *preference reversal*, in which participants price a low-probability high-payoff bet (e.g., a 3/36 chance to win \$100) above a high-probability low-payoff bet (e.g., a 28/36 chance to win \$10) even though they prefer the latter to the former when facing a simple choice between them (see, for example, Tversky *et al.*, 1990). Second, people tend to be more risk averse when matching prospects by varying probabilities than when matching prospects by varying outcomes (Hershey and Schoemaker, 1985). For instance, suppose that a participant is asked to report what p of receiving \$100 (or else nothing) is equally attractive to receiving \$35 for sure, and this participant reports a probability of .5. If that same participant is asked what certain amount is equally attractive to a .5 chance of \$100, he will generally report a value greater than \$35.

A popular alternative for overcoming limitations of direct matching procedures is to estimate cash equivalents from a series of choices. For instance, in pricing the prospect (100, .5) that offers a .5 chance of \$100, participants can be offered a series of choices between (100, .5) or \$100 for sure, (100, .5) or \$90 for sure, and so forth. For instance, if a participant

---

[9] This method is only incentive-compatible assuming the independence axiom, which of course is violated in prospect theory. For a further discussion see Karni and Safra, 1987.

chooses \$40 for sure over (100, .5) but also chooses (100, .5) over \$30 for sure, then by linear interpolation we can estimate his cash equivalent as approximately \$35. If a researcher tells participants that a randomly selected choice (from a randomly selected trial) will be honored for real money, then this method will be incentive-compatible (i.e., participants will have an economic incentive to respond honestly).

Sure amounts can be evenly spaced (e.g., Tversky and Fox, 1995) or logarithmically spaced (e.g., Tversky and Kahneman, 1992). If a researcher wishes to obtain higher-resolution estimates of cash equivalents, the sequential choice method cannot be readily accomplished in a single round. One approach is to use an iterated procedure in which a first-course evaluation is made followed by a more detailed series of choices etc. (e.g., Tversky and Kahneman, 1992; Tversky and Fox, 1995; Gonzalez and Wu, 1999). For instance, if a participant prefers \$40 to (100, .5) but \$30 to (100, .5) then four more choices might be presented between (100, .05) and \$28, \$26, \$24, \$22. Another, maximally efficient, approach is the "bisection method" in which each time a choice is made between two prospects (e.g. a risky and sure prospect) one of the outcomes is adjusted in smaller and smaller increments as preferences reverse. For instance, if a participant prefers \$50 to (100, .5) then he would be presented with a choice between \$25 and (100, .5). If he prefers the sure amount this time then he would be presented a choice between \$37.50 and (100, .5), and so forth. We note that, unlike single-round elicitations, the multi-round and bisection approaches to eliciting cash equivalents cannot easily be made incentive-compatible because if a randomly selected choice is honored for real money then participants can "game" the system so that a greater number of choices offer higher sure amounts (e.g., Harrison 1986). Pragmatically, however, this method remains popular, and there is no evidence that participants engage in such "gaming" (Peter Wakker, personal communication).

Empirical tests indicate that the bisection method performs much better than direct elicitation of cash equivalents (Bostic *et al.*, 1990). Fischer *et al.* (1999) noted that elicitation of cash equivalents through a series of choices will suffer from some of the problems of direct elicitation when the goal of determining cash equivalents is transparent. This can be obscured by eliciting choices in a staggered order so that each successive choice entails measurement of the cash equivalent of a different prospect. The downside to this approach is that it is more time consuming than a more straightforward application of the bisection or sequential choice method that prices one prospect at a time.

Modeling choice variability

The elicitation methods described thus far have all assumed a deterministic model of decision under risk. Naturally, one would not expect a decision maker's choices in practice to be 100% consistent. At different moments in time, a participant may reverse preferences between prospects. Such reversals may be due to decision errors (i.e., carelessness or lapses in concentration) and/or transitory variations in the participant's genuine underlying preferences (e.g., due to emotional, motivational, and cognitive states that influence risk preference). Reversals in preference are more likely to occur when the participant has difficulty distinguishing between prospects or has only weak preferences between them – if a decision maker is indifferent between prospects g_1 and g_2, then one would expect a 50% chance of reversing preferences on a subsequent choice between the prospects; the more strongly g_1 is preferred to g_2 the more often we expect it to be chosen. Such response variability is typically substantial in studies of risky choice. For instance, in a survey of eight studies of risky choice, Stott (2006, Table 11.1) found a median 23% rate of reversal in preferences when participants chose between the same pair of prospects on separate occasions within or across sessions.

There are two distinct approaches to modeling choice variability. The first is to assume that preferences are consistent with prospect theory but allow preferences consistent with that theory to vary from moment to moment. The "random preference" approach assumes that choices reflect a random draw from a probability distribution over preferences that are consistent with an underlying core theory (see Becker *et al.*, 1963, for an articulation of such a model under expected utility, and Loomes and Sugden, 1995, for a generalization). For instance, one could implement such a model using prospect theory value and weighting functions with variable parameters.

The second approach assumes a deterministic core theory but allows a specified error distribution to perturb the participant's response (see Becker *et al.*, 1963, for an application to EU). Formally, let $f(g_1, g_2)$ be the relative frequency with which prospect g_1 is selected over prospect g_2 in a pairwise choice. Decisions are assumed to be stochastically independent from one another and symmetric, so that $f(g_1, g_2) = 1 - f(g_2, g_1)$. Let $V(g_i)$ be the prospect theory value of prospect g_i. Most response variability models assume that $f(g_1, g_2)$ increases monotonically with $V(g_1) - V(g_2)$, the difference in prospect theory value of prospects 1 and 2.

The choice function $f(\cdot)$ can take several forms (see Stott, 2006, Table 11.4). First, it can manifest itself as a *constant* error function in which there is a fixed probability of expressing one's true preference. Thus, $f(g_1, g_2) = \varepsilon$ whenever $V(g_1) < V(g_2)$, ½ whenever $V(g_1) = V(g_2)$, $1 - \varepsilon$ whenever $V(g_1) > V(g_2)$, where $0 \leq \varepsilon \leq$ ½. Second, choice frequency might depend on the difference in prospect theory value between prospects, either following a *probit* transformation (e.g., Hey and Orme, 1994) or a *logit* transformation (e.g., Carbone and Hey, 1995). Thus, for the probit transformation,

$$f(g_1, g_2) = \Phi[(Vg_1) - V(g_2),\ 0, \sigma]$$

where $\Phi[x, \mu, \sigma]$ is the cumulative normal distribution with mean μ and SD σ at point x. Third, the choice function might follow a Luce (1959) choice rule, in which choice frequency depends on the ratio of prospect theory values of the prospects:

$$f(g_1, g_2) = \frac{V(g_1)^{\varepsilon}}{V(g_1)^{\varepsilon} + V(g_2)^{\varepsilon}}.$$

In an empirical test of several stochastic models assuming EU, Loomes and Sugden (1998) found that the random preference model tended to under-predict observed violations of dominance, and the error model assuming a probit transformation tended to over-predict such violations. The constant error form performed poorly.

The most comprehensive test to date of various choice functions and prospect theory value and weighting functional forms was reported by Stott (2006), who tested various combinations, including most of those described in this chapter. In his test, the model with the greatest explanatory power (adjusted for degrees of freedom) relied on a power value function (V1), a Prelec (1998) one-parameter weighting function (W3), and a logit function. However, for reasons already mentioned we recommend use of a two-parameter weighting function (W2) or (W3A).

The aforementioned models have been used to model preferences among pure gain or loss prospects. A stochastic method for measuring loss aversion was introduced by Tom *et al.* (2007). Their method required participants to make a series of choices as to whether or not to accept mixed prospects that offered a 50–50 chance of gaining \$$x$ or losing \$$y$ in which x and y were independently varied. These authors then assumed a piecewise linear value function, and also $w^{+}(.5) = w^{-}(.5)$[10]. They then determined the weight afforded the gain and loss portion of the gamble through logistic regression. This method has the advantage of allowing separate measurement of sensitivity to gains and losses (the regression coefficients), as well as response bias to accept or reject gambles (the intercept term).

NEUROSCIENTIFIC DATA

There has been substantial progress in understanding the neural correlates of prospect theory since we last reviewed the literature (Trepel *et al.*, 2005). Below, we first outline some challenges to effective characterization of the relation between neural activity and theoretical quantities, and then review recent work that has characterized the brain systems involved in various components of prospect theory.

Paradigmatic Challenges

Integrating theories from behavioral decision-making research with neuroscientific evidence has posed a number of challenges to researchers in both fields.

Developing Clean Comparisons

A neuroimaging study is only as good as its task design. In particular, in the context of behavioral decision theory it is critical that tasks cleanly manipulate particular theoretical quantities or components. For example, a study designed to examine the nature of probability weighting must ensure that the manipulation of probability does not also affect value. Because it is often impossible cleanly to isolate quantities in this way using any specific task, another alterative is to vary multiple quantities simultaneously and then model these manipulations parametrically. This allows the response to each quantity to be separately estimated. For example, Preuschoff *et al.* (2006) manipulated both expected reward and risk in a gambling task, and were able to demonstrate different regions showing parametric responses to each variable.

Isolating Task Components

One of the most difficult challenges of fMRI is the development of task paradigms and analytic approaches that allow isolation of specific task components. For example, in tasks where participants make a

---

[10] The former assumption is a customary and reasonable first approximation, and the latter assumption accords reasonably well with the data when it has been carefully tested (see Abdellaoui *et al.*, 2007c).

decision and then receive an outcome, it is desirable to be able separately to estimate the evoked response to the decision and to the outcome. Because the fMRI signal provides a delayed and smeared representation of the underlying neuronal activity, the evoked response lags the mental event by several seconds. A number of earlier studies used an approach where specific timepoints following a particular component are assigned to that component; however, this approach is not a reliable way to isolate trial components, as it will provide at best a weighted average of nearby events (Zarahn, 2000). It is possible to model the individual components using the general linear model, but the regressors that model the different components are often highly correlated, resulting in inflated variance. One solution to this problem involves the use of random-length intervals between trial components; this serves to decorrelate the model regressors for each task component and allow more robust estimation of these responses (see, for example, Aron *et al*., 2004).

Inferring Mental States from Neural Data

It is very common in the neuroeconomics literature to infer the engagement of particular mental states from neuroimaging data. For example, Greene *et al.* (2001) found that moral decision making for "personal" moral dilemmas was associated with greater activity in a number of regions associated with emotion (e.g., medial frontal gyrus) compared to "impersonal" moral dilemmas. On the basis of these results, they concluded that the difference between these tasks lies in the engagement of emotion when reasoning about the personal dilemmas. Poldrack (2006) referred to this approach as "reverse inference," and showed that its usefulness is limited by the selectivity of the activation in question. That is, if the specific regions in question only activate for the cognitive process of interest, then reverse inference may be relatively powerful; however, there is little evidence for strong selectivity in current neuroimaging studies, and this strategy should thus be used with caution. For example, ventral striatal activity is often taken to imply that the participant is experiencing reward, but activity in this region has also been found for aversive stimuli (Becerra *et al*., 2001) and novel non-rewarding stimuli (Berns *et al*. 1997), suggesting that this reverse inference is not well founded.

Reference-dependence and Framing Effects

The neural correlates of reference-dependence in decision making have been examined in two studies. De Martino *et al*. (2006) manipulated framing in a decision task in which participants chose between a sure outcome and a gamble after receiving an initial endowment on each trial; gambles were not resolved during scanning. Framing was manipulated by offering participants a choice between a sure loss and a gamble (e.g., lose £30 vs gamble) or a sure win and a gamble (e.g., keep £20 vs gamble). Participants showed the standard behavioral pattern of risk seeking in the loss frame and risk aversion in the gain frame, with substantial individual variability. Amygdala activity was associated with the dominant choices, with increased activity for sure choices in the gain frame and risky choices in the loss frame; the dorsal anterior cingulate cortex (ACC) showed an opposite pattern across conditions. Individual differences in behavioral framing-related bias were correlated with framing-related activation in orbitofrontal and medial prefrontal cortex; that is, participants who showed less framing bias (and thus "behaved more rationally") showed more activity for sure choices in the gain frame and risky choices in the loss frame compared to the other two conditions. Thus, whereas amygdala showed the framing-related pattern across all participants on average, in the orbitofrontal cortex (OFC) this pattern was seen increasingly for participants who showed less of a behavioral framing effect. Although amygdala activation is often associated with negative outcomes, it has also been associated with positive outcomes (e.g., Ghahremani and Poldrack, unpublished work; Weller *et al*., 2007), and the correlation of amygdala activity with choice in the de Martino study is consistent with coding of value in the amygdala.

Windmann *et al*. (2006) compared two versions of the Iowa Gambling Task (IGT): a "standard" version (in which participants must learn to choose smaller constant rewards in order to avoid large punishments) and an "inverted" version (in which participants must choose large constant punishments in order to obtain large rewards). This is similar to an inverted version of the IGT examined by Bechara *et al*. (2000), who found that patients with ventromedial prefrontal cortex (PFC) lesions were equally impaired on the standard and inverted versions of the task. Windmann *et al*. (2006) found that the inverted IGT was associated with a greater neural response to rewards compared to punishments in the lateral and ventromedial OFC when contrasted with the standard task. Interestingly, it appeared that some of the same lateral OFC regions activated for punishments vs rewards in the standard task were also activated for rewards vs punishments in the inverted task. These results suggest that the OFC response to outcomes is strongly modulated by the framing of outcomes. However, it is difficult to interpret results strongly from the IGT because of its conflation of risk and ambiguity. Because participants begin the task with no knowledge about the relevant probabilities and must learn them over time, it is not possible to know whether activation in the task reflects differences in the learning process or differences in the representation of value and/or probability.

Together, these studies provide initial evidence for the neural basis of framing effects, but much more work is needed. In particular, because neuroimaging methods are correlational, it is difficult to determine whether these results reflect the neural causes or neural effects of reference-dependence. Further work with lesion patients should provide greater clarity on this issue.

Value Function

Before reviewing papers that purport to examine neurophysiological correlates of the prospect theory value function, we pause to distinguish different varieties of utility. Traditionally, the utility construct in neoclassical economics refers to a hypothetical function that cannot be directly observed mapping states of wealth to numbers; a decision maker whose choices adhere to the four axioms reviewed in the first section of this chapter can be represented as maximizing expected utility. Thus, utility is a mathematical construct that may or may not reflect the mental states of decision makers.

Although prospect theory also has an axiomatic foundation (Wakker and Tversky, 1993), the model is motivated by behavioral phenomena, such as the psychophysics of diminishing sensitivity, that are assumed to correspond to mental states of decision makers. However, it is important to distinguish different varieties of utility when using tools of neuroscience to interpret mental states of decision makers. In particular, "utility" in the context of making a decision may not be the same thing as "utility" in the context of experiencing or anticipating the receipt of an outcome. Economists have focused primarily on a measure of what Kahneman *et al*. (1997) call *decision utility*, which is the weight of potential outcomes in decisions. However, as these authors point out, the original concepts of utility from Bentham and others focused on the immediate experience of pleasure and pain, which they refer to as *experienced utility*. Others have highlighted the importance of the utility related to anticipating a positive or negative outcome (e.g., Loewenstein, 1987), referred to as *anticipation utility*. Of particular interest is the fact that these different forms of utility can be dissociated; for example, individuals sometimes make decisions that serve to decrease their experienced or anticipation utility. In order to be able to interpret clearly the results of neuroimaging studies, it is critical to distinguish between these different forms of utility.

The distinction between different forms of utility in behavioral decision theory parallels the distinction between "wanting" and "liking" that has developed in the animal literature (Berridge, 2007). A large body of work has shown that the neural systems involved in motivating aspects of reward ("wanting") can be dissociated from those involved in the hedonic aspects of reward ("liking"). This work has largely focused on neurochemical dissociations. Whereas dopamine is often thought to be involved with pleasurable aspects of reward, a large body of work in rodents has shown that disruption of the dopamine system impairs animals' motivation to obtain rewards (particularly when effort is required), but does not impair their hedonic experience (as measured using conserved behavioral signals of pleasure such as tongue protrusion and paw licking; Pecina *et al*., 2006). The hedonic aspects of reward appear to be mediated by opioid systems in the ventral striatum and pallidum. Although the mapping of neurochemical systems to functional neuroimaging results is tricky (Knutson and Cooper, 2005), these results provide further suggestion that "utility" is not a unitary concept.

Because it is most directly relevant to the prospect theory value function, we focus here on decision utility. This is the value signal that is most directly involved in making choices, particularly when there is no immediate outcome of the decision, as in purchasing a stock or lottery ticket. It has received relatively little interest in the neuroeconomics literature compared to experienced and anticipation utility, but several recent studies have examined the neural basis of decision utility using fMRI. Tom *et al*. (2007) imaged participants during a gamble acceptability paradigm, in which participants decided whether to accept or reject mixed gambles offering a 50% chance of gain and 50% chance of loss. The size of the gain and loss were varied parametrically across trials, with gains ranging from \$10 to \$40 (in \$2 increments) and losses from \$5 to \$20 (in \$1 increments). Participants received an endowment in a separate session 1 week before scanning, in order to encourage integration of the endowment into their assets and prevent the risk-seeking associated with "house money" effects (Thaler and Johnson, 1990). Participants exhibited loss-averse decision behavior, with a median loss aversion parameter $\lambda = 1.93$ (range: 0.99 to 6.75). Parametric analyses examined activation in relation to gain and loss magnitude. A network of regions (including ventral and dorsal striatum, ventromedial and ventrolateral

PFC, ACC, and dopaminergic midbrain regions) showed increasing activity as potential gain increased. Strikingly, no regions showed increasing activity as potential loss increased (even using weak thresholds in targeted regions including amygdala and insula). Instead, a number of regions showed *decreasing* activation as losses increased, and these regions overlapped with the regions whose activity increased for increasing gains.

The Tom *et al.* (2007) study further characterized the neural basis of loss aversion by first showing that a number of regions (including ventral striatum) showed "neural loss aversion," meaning that the decrease in activity for losses was steeper than the increase in activity for gains. Using whole-brain maps of these neural loss aversion parameters, they found that behavioral loss aversion was highly correlated across individuals with neural loss aversion in a number of regions including ventral striatum and ventrolateral PFC. These data are strongly consistent with prospect theory's proposal of a value function with a steeper slope for losses than for gains.

Decision utility was examined by Plassmann *et al.* (2007) using a "willingness-to-pay" (WTP) paradigm in which participants placed bids for a number of ordinary food items in a Becker–DeGroot–Marschak (BDM) auction, which ensures that participants' choices are an accurate reflection of their preferences. "Free bid" trials, in which participants decided how much to bid on the item, were compared with "forced bid" trials, in which participants were told how much to bid. Activity in ventromedial and dorsolateral PFC was correlated with WTP in the free bid trials but not the forced bid trials, suggesting that these regions are particularly involved in coding for decision utility.

The neural correlates of purchasing decisions were also examined by Knutson *et al.* (2007). Participants were presented at each trial with a product, and then given a price for that product and asked to indicate whether they would purchase the product for that price. Participants also provided WTP ratings after scanning was completed. Activity in ventral striatum and ventromedial PFC was greater for items that were purchased, whereas activity in anterior insula was lower for items that were purchased. A logistic regression analysis examined whether decisions could be better predicted by self report data or brain activity; although self-report data were much more predictive of purchasing decisions, a small (~1% of variance) increase in predictability was obtained when self-report and fMRI data were combined.

Because of the oft-noted association of the amygdala with negative emotions, it might be suspected that it would be involved in loss aversion in decision making. However, only one study has found amygdala activity in relation to loss aversion. Weber *et al.* (2007) examined reference-dependence using a design in which participants either bought or sold MP3 songs in a BDM auction. Comparison of selling trials versus buying trials showed greater activity in both amygdala and dorsal striatum, whereas comparison of buying versus selling trials showed greater activity in the parahippocampal gyrus. Given the association of amygdala with both positive and negative outcomes, it is unclear whether the effect for selling versus buying reflects the disutility of losing a good, the utility of gaining money, or some other factor. Further, a recent study by Weller *et al.* (2007) shows that patients with amygdala damage are actually impaired in making decisions about potential gains, whereas they are unimpaired in decisions about potential losses. These findings highlight the complexity of the amygdala's role in decision making, potentially suggesting that there are underlying factors modulating amygdala activity that have yet to be discovered.

Together, these results begin to characterize a system for decision utility, with the ventromedial PFC appearing as the most consistent region associated with decision utility. These results are consistent with other data from neurophysiology in non-human primates suggesting a representation of the value of goods such as foods (Padoa-Schioppa and Assad, 2006). However, the results also raise a number of questions. First, they cast some doubt over a simple two-system model with separate regions processing potential gains and losses. It is clear that the neural activity evoked by potential gains and losses is only partially overlapping with that evoked by actual gains and losses, but further work is needed to better characterize exactly how the nature of the task (such as the participants' anticipation of outcomes) changes neural activity. Second, they cast doubt over the common inference that amygdala activity is related to negative emotion, as it is clear that positive outcomes can also activate the amygdala. Further work is necessary to better understand the amygdala's role in decision making. Third, they leave unexplained how neural activity relates to the characteristic S-shaped curvature of the value function that contributes a tendency toward risk aversion for gains and risk seeking for losses.

Probability Weighting Distortions

A number of recent studies have attempted to identify neural correlates of distortions in probability weighting. Paulus and Frank (2006) used a certainty equivalent paradigm in which participants chose between a gamble and a sure outcome on each trial; the gamble was altered in successive trials to estimate the certainty equivalent. Non-linearity of the probability weighting function was estimated using the Prelec (1998) weighting function. Regression of activation for high- versus low-probability prospects showed that activity in the ACC was correlated with the non-linearity parameter, such that participants with more ACC activity for high versus low prospects were associated with more linear weighting of probabilities.

Non-linearities in probability weighting were also examined by Hsu *et al*. (2008). Participants chose between pairs of simple gambles, which varied in outcome magnitude and probability; on each trial, each gamble was first presented individually, then they were presented together and the participant chose between them. Weighting function non-linearity was estimated using the Prelec (1998) one-parameter weighting function (W3B). In order to isolate regions exhibiting non-linear responses with probability, separate regressors were created which modeled a linear response with *p* and a deflection from that linear function which represents non-linear effects. Significant correlations with both linear and non-linear regressors were found in several regions, including the dorsal striatum. Further analysis of individual differences showed a significant correlation between behavioral non-linearity and non-linearity of striatal response across participants.

Probability weighting distortion for aversive outcomes was examined by Berns *et al*. (2007). In a first phase, participants passively viewed prospects which specified the magnitude and probability of an electric shock. In a second phase, participants chose between pairs of lotteries. A quantity was estimated ("neurological probability response ratio," NPRR) which indexed the response to a lottery with probability less than one to a lottery with a probability of one (normalized by respect to the response to probability 1/3, which is the sampled point nearest to the likely intersection of the non-linear weighting function and linear weighting function – see Figure 11.3e). For the passive phase, NPRR was significantly non-linear for most regions examined, including regions in the dorsal striatum, prefrontal cortex, insula, and ACC. Activity from the passive phase was also used to predict choices during the choice phase; the fMRI signals provided significant predictive power, particularly for lotteries that were near the indifference point. Thus, there appears to be fairly wide-scale overweighting of low-probability aversive events in a number of brain regions.

Although the results of these studies are preliminary and not completely consistent, they suggest that it should be possible to identify the neural correlates of probability weighting distortions. It will be important to determine which regions are causally involved in these distortions (as opposed to simply reflecting the distortions) by testing participants with brain lesions or disorders. If non-linearities are the product of a specific brain system, then it should be possible to find participants whose choices are rendered linear with probability following specific lesions, similar to findings that VMPFC lesions result in more advantageous behavior in risky choice (Shiv *et al*., 2005).

CONCLUSIONS AND FUTURE DIRECTIONS

The field of neuroeconomics is providing a rapidly increasing amount of data regarding the phenomena that lie at the heart of prospect theory, such as framing effects and loss aversion. But we might ask: what have these data told us about prospect theory? It is clear that the demonstrations of neural correlates of several of the fundamental behavioral phenomena underlying prospect theory (loss aversion, framing effects, and probability weighting distortions) provide strong evidence to even the most entrenched rational choice theorists that these "anomalies" are real. The data have also started to provide more direct evidence regarding specific claims of the theory.

Our review of behavioral and neuroscience work on prospect theory and the neuroscience of behavioral decision making suggests a number of points of caution for future studies of decision making in the brain:

1. It is critical to distinguish between the different varieties of utility in designing and interpreting neuroscience studies. Studies in which participants make a decision and then receive an immediate outcome may be unable to disentangle the complex combination of decision, anticipation, and experienced utilities that are likely to be in play in such a task.
2. Under prospect theory, risk attitudes toward different kinds of prospects are interpreted in different ways. Risk aversion for mixed gambles is attributed to loss aversion; the fourfold pattern of risk attitudes for pure gain or loss gambles is attributed to diminishing sensitivity both to money (as reflected by curvature of the value function) and probability (as reflected by the inverse S-shaped weighting function). It is easy to conflate these factors empirically; for instance, if one assumes a single-parameter weighting function that only allows variation in curvature

but not elevation, then variations in observed risk attitudes across all probability levels may be misattributed to curvature of the value function.

3. Reverse inference (i.e., the inference of mental states from brain-imaging data) should be used with extreme care. As a means for generating hypotheses it can be very useful, but its severe limitations should be recognized.

Challenges for the Future

As neuroeconomics charges forward, we see a number of important challenges for our understanding of the neurobiology of prospect theory. First, it is critical that neuroimaging studies are integrated with studies of neuropsychological patients in order to determine not just which regions are correlated with particular theoretical phenomena, but also whether those regions are necessary for the presence of the phenomena. A nice example of this combined approach was seen in the study of ambiguity aversion by Hsu *et al.* (2005). It is likely that many of the regions whose activity is correlated with theoretical quantities (e.g., curvature of weighting function) may be effects rather than causes of the behavioral phenomena.

Another challenge comes in understanding the function of complex neural structures, such as the ventral striatum and amygdala, in decision making. Each of these regions is physiologically heterogeneous, but the resolution of current imaging techniques leads them to be treated as singular entities. In the amygdala, the heterogeneous nuclei are large enough that they could potentially be differentiated using currently available neuroimaging methods (e.g., Etkin *et al.*, 2004). The neurobiological heterogeneity of the ventral striatum is more difficult to address using current neuroimaging methods; there are both structural features that are not currently visible to human neuroimaging (e.g., accumbens core vs. shell) as well as substantial cellular heterogeneity (e.g., striosomes vs. matrix, direct vs. indirect pathway) at an even finer grain. Finally, there is still substantial controversy over the degree to which imaging signals in the ventral striatum reflect dopamine release as opposed to excitatory inputs or interneuron activity. It is clear that imaging signals in the ventral striatum often exhibit activity that parallels the known patterns of dopamine neuron firing (in particular, prediction error signals), and dopamine has strong vascular as well as neuronal effects, so it is likely that it exerts powerful effects on imaging signals, but it is not currently known how to disentangle these effects from local neuronal effects.

Finally, one critical extension of present work will be to relate it to other work in the domain of cognitive control. The role of frontal and basal ganglia regions in the control of cognitive processes (including inhibition, selection, and interference resolution) is becoming increasingly well specified, but how these processes relate to decision making remains unknown. Given the availability of the prefrontal cortex to both neuroimaging and disruption by transcranial magnetic stimulation (TMS), there is hope that an understanding of the relation between cognitive control and decision making will be relatively tractable in comparison to subcortical regions.

APPENDIX

Formal Presentation of Cumulative Prospect Theory (adapted from Tversky and Kahneman, 1992)

Let S be the set whose elements are interpreted as states of the world, with subsets of S called *events*. Thus, S is the certain event and ϕ is the null event. A weighting function W (on S), also called a *capacity*, is a mapping that assigns to each event in S a number between 0 and 1 such that $W(\phi) = 0$, $W(S) = 1$, and $W(A) \geq W(B)$ if and only if $A \supseteq B$.

Let X be a set of consequences, also called *outcomes*, that also includes a neutral outcome 0. An uncertain prospect f is a function from S into X that assigns to each event A_i a consequence x_i. Assume that the consequences are ordered by magnitude so that $x_i > x_j$ if $i > j$. Cumulative prospect theory separates prospects into a positive part, f^+, that includes all $x_i > 0$, and a negative part, f^-, that includes all $x_i < 0$. CPT assumes a strictly increasing value function $v(x)$ satisfying $v(x_0) = v(0) = 0$.

CPT assigns to each prospect f a number $V(f)$ such that $f \succeq g$ if and only if $V(f) \geq V(g)$. Consider a prospect $f = (x_i, A_i)$, $-m \leq i \leq n$, in which positive (negative) subscripts refer to positive (negative) outcomes and decision weights $\pi^+(f^+) = (\pi_0^+, \ldots, \pi_n^+)$ and $\pi^-(f^-) = (\pi_{-m}^-, \ldots, \pi_0^-)$ for gains and losses, respectively. The value V of the prospect is given by

$$V(f) = V(f^+) + V(f^-)$$

where

$$V(f^+) = \sum_{i=1}^{n} \pi_i^+ v(x_i), \quad \text{and} \quad V(f^-) = \sum_{i=-m}^{0} \pi_i^- v(x_i)$$

where π^+ and π^- are defined as follows:

$$\pi_n^+ = W^+(A_n), \quad \pi_{-m}^- = W^-(A_{-m})$$

$$\pi_i^+ = W^+(A_i \cup \ldots \cup A_n) - W^+(A_{i+1} \cup \ldots \cup A_n), \text{ for } 0 \leq i \leq n-1$$

$$\pi_i^- = W^-(A_{-m} \cup \ldots \cup A_i) - W^-(A_{-m} \cup \ldots \cup A_{i-1}), \text{ for } 1-m \leq i \leq 0.$$

Acknowledgments

We thank Mohammed Abdellauoi, Han Bleichrodt, Paul Glimcher and Peter Wakker for useful feedback on earlier versions of this chapter and Liat Hadar for helpful assistance.

References

Abdellaoui, M. (2000). Parameter-free elicitation of utility and probability weighting functions. *Management Sci.* 46, 1497–1512.

Abdellaoui, M., Vossmann, F., and Weber, M. (2005). Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. *Management Sci.* 51, 1384–1399.

Abdellaoui, M., Barrios, C., and Wakker, P.P. (2007a). Reconciling introspective utility with revealed preference: experimental arguments based on prospect theory. *J. Econometrics* 138, 356–378.

Abdellaoui, M., Bleichrodt, H., and Paraschiv, C. (2007b). Measuring loss aversion under prospect thoery: A parameter-free approach. *Management Sci.* 53, 1659–1674.

Abdellaoui, M., Bleichrodt, H., L'Haridon, O. (2007c). A tractable method to measure utility and loss aversion under prospect theory. Unpublished manuscript, HEC, April.

Allais, M. (1953). Le comportement de l'homme rationel devant le risque, critique des postulates et axiomes de l'école americaine. *Econometrica* 21, 503–546.

Allais, M. and Hagen, O. (1979). The so-called Allais paradox and rational decisions under uncertainty. In: O.H.M. Allais (ed.), *Expected Utility Hypothesis and the Allais Paradox*. Dordrecht: Reidel Publishing Company, pp. 434–698.

Aron, A.R., Monsell, S., Sahakian, B.J., and Robbins, T.W. (2004). A componential analysis of task-switching deficits associated with lesions of left and right frontal cortex. *Brain* 127, 1561–1573.

Aron, A.R., Shohamy, D., Clark, J. *et al*. (2004). Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *J. Neurophysiol.* 92, 1144–1152.

Barberis, N. and Xiong, W. (2006). *What Drives the Disposition Effect? An Analysis of a Long-standing Preference-based Explanation*. Cambridge, MA: National Bureau of Economic Research.

Barberis, N., Huang, M., and Santos, T. (2001). Prospect theory and asset prices. *Q. J. Economics* 116, 1–53.

Bateman, I., Munro, A., Rhodes, B. *et al*. (1997). A test of the theory of reference-dependent preferences. *Q. J. Economics*, 112, 470–505.

Becerra, L., Breiter, H.C., Wise, R. *et al*. (2001). Reward circuitry activation by noxious thermal stimuli. *Neuron* 32, 927–946.

Bechara, A., Tranel, D., and Damasio, H. (2000). Characterization of the decision making deficit of patients with ventromedial prefrontal cortex lesions. *Brain* 123, 2189–2202.

Becker, G.M., DeGroot, M.H., and Marschak, J. (1963). Stochastic models of choice behavior. *Behavioral Sci.* 8, 41–55.

Becker, G.M., DeGroot, M.H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Sci.* 9, 226–232.

Benartzi, S. and Thaler, R.H. (1995). Myopic loss aversion and the equity premium puzzle. *Q. J. Economics* 110, 73–92.

Bernoulli, D. (1954/1738). Exposition of a new theory on the measurement of risk [translation by L. Sommer of D. Bernoulli, 1738, Specimen theoriae novae de mensura sortis, *Papers of the Imperial Academy of Science of Saint Peterburg* 5, 175–192]. *Econometrica* 22(1), 23–36.

Berns, G.S., Cohen, J.D., and Mintun, M.A. (1997). Brain regions responsive to novelty in the absence of awareness. *Science* 276, 1272–1275.

Berns, G.S., Capra, C.M., Chappelow, J. *et al*. (2007). Nonlinear neurobiological probability weighting functions for aversive outcomes. *NeuroImage* 39, 2047–2057.

Berridge, K.C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacol. (Berl.)* 191, 391–431.

Bleichrodt, H. and Pinto, J.L. (2000). A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management Sci.* 46, 1485–1496.

Bostic, R., Herrnstein, R., and Luce, R.D. (1990). The effect on the preference-reversal phenomenon of using choice indifferences. *J. Econ. Behav. Org.* 13, 192–212.

Camerer, C.F. and Ho, T.H. (1994). Violations of the betweenness axiom and nonlinearity in probability. *J. Risk Uncertainty* 8, 167–196.

Camerer, C.F. and Hogarth, R.M. (1999). The effects of financial incentives in experiments: a review and capital-labor-production framework. *J. Risk Uncertainty* 19, 7–42.

Camerer, C.F. and Weber, M. (1992). Recent developments in modeling preferences: uncertainty and ambiguity. *J Risk Uncertainty* 5, 325–370.

Camerer, C., Babcock, L., Loewenstein, G., and Thaler, R. (1997). Labor supply of New York city cab drivers: one day at a time. *Q. J. Economics* 111, 408–441.

Carbone, E. and Hey, J.D. (1994). Which error story is best? *J. Risk Uncertainty* 20, 161–176.

Chow, C.C. and Sarin, R.K. (2001). Comparative ignorance and the Ellsberg paradox. *J. Risk Uncertainty* 22, 129–139.

Cook, P.J. and Clotfelter, C.T. (1993). The peculiar scale economies of Lotto. *Am. Econ. Rev.* 83, 634–643.

De Martino, B., Kumaran, D., Seymour, B., and Dolan, R.J. (2006). Frames, biases, and rational decision-making in the human brain. *Science* 313, 684–687.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *Q. J. Economic*, 75, 643–669.

Etchart-Vincent, N. (2004). Is probability weighting sensitive to the magnitude of consequences? An experimental investigation on losses. *J. Risk Uncertainty*, 28, 217–235.

Etkin, A., Klemenhagen, K.C., Dudman, J.T. *et al*. (2004). Individual differences in trait anxiety predict the response of the basolateral amygdala to unconsciously processed fearful faces. *Neuron* 44, 1043–1055.

Fehr, E. and Gotte, L. (2007). Do workers work more if wages are high? Evidence from a randomized field experiment. *Am. Econ. Rev.* 97, 298–317.

Fehr-Duda, H., Bruin, A., Epper, T. F., and Schubert, R. (2007). Rationality on the rise: why relative risk aversion increases with stake size. Working Paper #0708, University of Zurich.

Fennema, H. and Van Assen, M. (1999). Measuring the utility of losses by means of the tradeoff method. *J. Risk Uncertainty*, 17, 277–295.

Fennema, H. and Wakker, P. (1997). Original and cumulative prospect theory: a discussion of empirical differences. *J. Behav. Decision Making* 10, 53–64.

Fischer, G.W., Carmon, Z., Ariely, D., and Zauberman, G. (1999). Goal-based construction of preferences: task goals and the prominence effect. *Management Sci.* 45, 1057–1075.

Fishburn, P. and Kochenberger, G. (1979). Two-piece von Neumann–Morgenstern utility functions. *Decision Sci.* 10, 503–518.

Fox, C.R. and Hadar, L. (2006). "Decisions from experience" = sampling error + prospect theory: reconsidering Hertwig, Barron, Weber and Erev (2004). *Judgment Decision Making* 1, 159–161.

Fox, C.R. and See, K.S. (2003). Belief and preference in decision under uncertainty. In: D. Hardman and L. Macchi (eds), *Reasoning and Decision Making: Current Trends and Perspectives*. New York, NY: Wiley, pp. 273–314.

Fox, C.R. and Tversky, A. (1995). Ambiguity aversion and comparative ignorance. *Q. J. Economics,* 110, 585–603.

Fox, C.R. and Tversky, A. (1998). A belief-based account of decision under uncertainty. *Management Sci.* 44, 879–895.

Fox, C.R. and Weber, M. (2002). Ambiguity aversion, comparative ignorance, and decision context. *Org. Behav. Hum. Dec. Proc.* 88, 476–498.

French, K.R. and Poterba, J.M. (1991). Investor diversification and international equity markets. *Am. Econ. Rev.* 81, 222–226.

Goldstein, W.M. and Einhorn, H.J. (1987). Expression theory and the preference reversal phenomenon. *Psychological Rev.* 94, 236–254.

Gonzalez, R. and Wu, G. (1999). On the shape of the probability weighting function. *Cogn. Psychol.* 38, 129–166.

Gonzalez, R. and Wu, G. (2003). Composition rules in original and cumulative prospect theory. Unpublished Paper.

Greene, J.D., Sommerville, R.B., Nystrom, L.E. *et al*. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105–2108.

Hadar, L. and Fox, C.R. (2008). *Deconstructing Uncertainty: The Impact of Experience, Belief, and Preference on Decisions*. Working Paper, UCLA.

Hardie, B.G.S., Johnson, E.J., and Fader, P.S. (1993). Modeling loss aversion and reference dependence effects on brand choice. *Marketing Sci.* 12, 378–394.

Harless, D.W. and Camerer, C.F. (1994). The predictive utility of generalized expected utility theories. *Econometrica* 62, 1251–1290.

Harrison, G.W. (1986). An experimental test for risk aversion. *Economic letters* 21, 7–11.

Heath, C. and Tversky, A. (1991). Preference and belief: ambiguity and competence in choice under uncertainty. *J. Risk Uncertainty* 4, 5–28.

Heath, C., Larrick, R.P., and Wu, G. (1999). Goals as reference points. *Cogn. Psychol.* 38, 79–109.

Hershey, J.C. and Schoemaker, P.J.H. (1980). Prospect theory's reflection hypothesis: a critical examination. *Org. Behav. Hum. Dec. Proc.* 25, 395–418.

Hershey, J.C. and Schoemaker, P.J.H. (1985). Probability versus certainty equivalence methods in utility measurement: are they equivalent? *Management Sci.* 31, 1213–1231.

Hertwig, R. and Ortmann, A. (2001). Experimental practices in economics: a methodological challenge for psychologists? *Behav. Brain Sci.* 24, 383–451.

Hertwig, R., Barron, G., Weber, E.U., and Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Sci.* 15, 534–539.

Hey, J.D. and Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica* 62, 1291–1326.

Holt, C.A. and Laury, S.K. (2002). Risk aversion and incentive effects. *Am. Econ. Rev.* 92, 1644–1655.

Horowitz, J.K. and McConnell, K.E. (2002). A review of WTA/WTP studies. *J. Environ. Econ. Management,* 44, 426–447.

Hsu, M., Bhatt, M., Adolphs, R. *et al*. (2005). Neural systems responding to degrees of uncertainty in human decision making. *Science* 310, 1680–1683.

Hsu, M., Zhao, C., and Camerer, C.F. (2008). *Neural Evidence for Nonlinear Probabilities in Risky Choice*. Working Paper, California Institute of Technology.

Jensen, N.E. (1967). An introduction to Bernoullian Utility Theory, I. *Swedish J. Economics* 69, 163–183.

Johnson, E.J. and Goldstein, D. (2003). Do defaults save lives? *Science* 302, 1338–1339.

Johnson, E.J., Gächter, S., and Herrmann, A. (2007). *Exploring the Nature of Loss Aversion*. IZA Discussion Paper, No. 2015.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 4, 263–291.

Kahneman, D. and Tversky, A. (1991). Loss aversion in riskless choice: a reference-dependent model. *Q. J. Economics* 106, 1039–1061.

Kahneman, D., Knetch, J.L., and Thaler, R.H. (1986). Fairness as a constraint on profit seeking: entitlements in markets. *Am. Econ. Rev.* 76, 728–741.

Kahneman, D., Knetsch, J.L., and Thaler, R.H. (1990). Experimental tests of the endowment effect and the Coase theorem. *J. Political Econ.* 98, 1325–1348.

Kahneman, D., Wakker, P.P., and Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *Q. J. Economics* 112, 375–405.

Karni, E. and Safra, Z. (1987). "Preference reversal" and the observability of preferences by experimental methods. *Econometrica* 55, 675–685.

Kilka, M. and Weber, M. (2001). What determines the shape of the probability weighting function under uncertainty? *Management Sci.* 47, 1712–1726.

Knetsch, J.L. (1989). The endowment effect and evidence of nonreversible indifference curves. *Economic Rev.* 79, 1277–1284.

Knight, F. (1921). *Risk, Uncertainty, and Profit*. Boston, MA: Houghton-Mifflin.

Knutson, B. and Cooper, J.C. (2005). Functional magnetic resonance imaging of reward prediction. *Curr. Opin. Neurol.* 18, 411–417.

Knutson, B., Rick, S., Wimmer, G.E. *et al*. (2007). Neural predictors of purchases. *Neuron* 53, 147–156.

Lattimore, P.K., Baker, J.R., and Witte, A.D. (1992). The influence of probability on risky choice – a parametric examination. *J. Econ. Behav. Org.* 17, 377–400.

Linville, P.W. and Fischer, G.W. (1991). Preferences for separating of combining events. *J. Pers. Social Psychol.* 60, 5–23.

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *Economic J.* 97, 666–684.

Loomes, G. and Sugden, G. (1995). Incorporating a stochastic element into decision theories. *Eur. Econ. Rev.* 39, 641–648.

Loomes, G. and Sugden, G. (1998). Testing different stochastic specifications of risky choice. *Economica* 65, 581–598.

Luce, R.D. (1959). *Individual Choice Behavior*. New York, NY: Wiley.

Luce, R.D. and Fishburn, P.C. (1991). Rank- and sign-dependent linear utility models for finite first-order gambles. *J. Risk Uncertainty* 4, 29–59.

March, J.G. and Shapira, Z. (1987). Managerial perspectives on risk and risk-taking. *Management Sci.* 33, 1404–1418.

McNeil, B.J., Pauker, S.G., Sox, H.C., Jr, and Tversky, A. (1982). On the elicitation of preferences for alternative therapies. *New Engl. J. Med.* 306, 1259–1262.

Odean, T. (1998). Are investors reluctant to realize their losses? *J. Finance* 53, 1775–1798.

Padoa-Schioppa, C. and Assad, J.A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.

Paulus, M.P. and Frank, L.R. (2006). Anterior cingulate activity modulates nonlinear decision weight function of uncertain prospects. *Neuroimage* 30, 668–677.

Payne, J.W., Laughhunn, D.J., and Crum, R. (1981). Further tests of aspiration level effects in risky choice. *Management Sci.* 27, 953–958.

Pecina, S., Smith, K.S., and Berridge, K.C. (2006). Hedonic hot spots in the brain. *Neuroscientist* 12, 500–511.

Plassmann, H., O'Doherty, J., and Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27, 9984–9988.

Poldrack, R.A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10, 59–63.

Prelec, D. (1998). The probability weighting function. *Econometrica* 66, 497–527.

Prelec, D. (2000). Compound invariant weighting functions in prospect theory. In: D. Kahneman and A. Tversky (eds), *Choices, Values, and Frames*. Cambridge: Cambridge University Press, pp. 67–92.

Preuschoff, K., Bossaerts, P., and Quartz, S.R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51, 381–390.

Rabin, M. (2000). Risk aversion and expected-utility theory: a calibration theorem. *Econometrica* 68, 1281–1292.

Rottenstreich, Y. and Tversky, A. (1997). Unpacking, repacking, and anchoring: advances in support theory. *Psychological Rev.* 2, 406–415.

Samuelson, W. and Zeckhauser, R. (1988). Status quo bias in decision making. *J. Risk Uncertainty* 1, 7–59.

Savage, L.J. (1954). *The Foundations of Statistics*. New York, NY: Wiley.

Shiv, B., Loewenstein, G., and Bechara, A. (2005). The dark side of emotion in decision making: when individuals with decreased emotional reactions make more advantageous decisions. *Cogn. Brain Res.* 23, 85–92.

Slovic, P. (1987). Perception of risk. *Science* 236, 280–285.

Stott, H.P. (2006). Cumulative prospect theory's functional menagerie. *J. Risk Uncertainty* 32, 101–130.

Thaler, R. (1980). Toward a positive theory of consumer choice. *J. Econ. Behav. Org.* 1, 39–60.

Thaler, R.H. (1985). Mental accounting and consumer choice. *Marketing Sci.* 4, 199–214.

Thaler, R.H. (1999). Mental accounting matters. *J. Behav. Decision Making,* 12, 183–206.

Thaler, R.H. and Johnson, E.J. (1990). Gambling with the house money and trying to break even: the effects of prior outcomes on risky choice. *Management Sci.* 36, 643–660.

Tom, S.M., Fox, C.R., Trepel, C., and Poldrack, R.A. (2007). The neural basis of loss aversion in decision making under risk. *Science* 315, 515–518.

Trepel, C., Fox, C.R., and Poldrack, R.A. (2005). Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk. *Brain Res. Cogn. Brain Res.* 23, 34–50.

Tversky, A. (1967). Additivity, utility, and subjective probability. *J. Math. Psychol.* 4, 175–201.

Tversky, A. and Fox, C.R. (1995). Weighing risk and uncertainty. *Psychological Rev.* 102, 269–283.

Tversky, A. and Kahneman, D. (1986). Rational choice and the framing of decisions. *J. Business,* 59(2), S251–S278.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory – cumulative representation of uncertainty. *J. Risk Uncertainty,* 5, 297–323.

Tversky, A. and Koehler, D.J. (1994). Support theory: a nonextensional representation of subjective probability. *Psychological Rev.* 101, 547–567.

Tversky, A. and Wakker, P. (1995). Risk attitudes and decision weights. *Econometrica* 63, 1255–1280.

Tversky, A., Slovic, P., and Kahneman, D. (1990). The causes of preference reversal. *Am. Econ. Rev.* 80, 204–217.

van de Kuilen, G., Wakker, P.P., and Zou, L. (2006). A midpoint technique for easily measuring prospect theory's probability weighting. Working Paper, Econometric Institute, Erasmus University, Rotterdam.

von Neumann, J. and Morgenstern, O. (1947). *Theory of Games and Economic Behavior*, 2nd edn. Princeton, NJ: Princeton University Press.

Wakker, P.P. (2001). Testing and characterizing properties of nonadditive measures through violations of the sure-thing principle. *Econometrica* 69, 1039–1059.

Wakker, P.P. (2004). On the composition of risk preference and belief. *Psychological Rev.* 111, 236–241.

Wakker, P. and Deneffe, D. (1996). Eliciting von Neumann–Morgenstern utilities when probabilities are distorted or unknown. *Management Sci.* 42, 1131–1150.

Wakker, P. and Tversky, A. (1993). An axiomatization of cumulative prospect theory. *J. Risk Uncertainty* 7, 147–176.

Wakker, P., Thaler, R., and Tversky, A. (1997). Probabilistic insurance. *J. Risk Uncertainty* 15, 7–28.

Weber, B., Aholt, A., Neuhaus, C. *et al.* (2007). Neural evidence for reference-dependence in real-market-transactions. *Neuroimage* 35, 441–447.

Weller, J.A., Levin, I.P., Shiv, B., and Bechara, A. (2007). Neural correlates of adaptive decision making for risky gains and losses. *Psychological Sci.* 18, 958–964.

Windmann, S., Kirsch, P., Mier, D. *et al.* (2006). On framing effects in decision making: linking lateral versus medial orbitofrontal cortex activation to choice outcome processing. *J Cogn. Neurosci.* 18, 1198–1211.

Wu, G. and Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Sci.* 42, 1676–1690.

Wu, G. and Gonzalez, R. (1998). Common consequence conditions in decision making under risk. *J. Risk Uncertainty* 16, 115–139.

Wu, G. and Gonzalez, R. (1999). Nonlinear decision weights in choice under uncertainty. *Management Sci.* 45, 74–85.

Wu, G. and Markle, A.B. (2008). An empirical test of gain-loss separability in prospect theory. *Management Sci.* (forthcoming).

Zarahn, E. (2000). Testing for neural responses during temporal components of trials with BOLD fMRI. *Neuroimage* 11, 783–796.

CHAPTER

12

Values and Actions in Aversion

Peter Dayan and Ben Seymour

OUTLINE

INTRODUCTION

It was the English scholar Jeremy Bentham who first argued that the understanding of human economic behavior might benefit from the study of the physiological processes from which it derives (Bentham, 1823). Pertinently, he pursued an account of economic decision making that balanced the opposing motives of losses and gains, in recognition of the fact that most choices involve contemplation of comparable measures of each. Furthermore, he recognized that the immutable characteristic of the former (incarnate as his plethora of "pains" (Bentham, 1817) is the basic devaluing property that drives decisions to reduce or avoid them.

More prosaically, issues of loss are central to many everyday economic decisions, such as health, insurance, and borrowing; further, apparent anomalies of choice such as loss aversion, framing effects, and regret all arise in aversive contexts. There is even a tight, though confusing, link between aversion and stress and psychiatric conditions such as depression. Nevertheless, partly for ethical reasons having to do with the undesirability of actually relieving human subjects of part of their own wealth in an experiment, it has been very hard to study truly aversive learning and processing in a human economic context. Fortunately, along with a number of inventive attempts along these lines, substantial data relevant to these issues have been collected in experimental psychology and behavioral neuroscience using other forms of aversive outcomes, and this chapter is underpinned by these results, along with the much more substantial understanding of reward, which is aversion's evil twin.

Through such sources, the broad outline of the architecture of decision making is slowly emerging. There is ample evidence that a number of systems is involved in making, and learning to make, predictions about future positive (which, in the psychological literature, are often called *appetitive*) and negative (*aversive*) outcomes, and in choosing actions that generally increase the former and decrease the latter (Adams and Dickinson, 1981; Dickinson and Balleine, 2002; Daw *et al*., 2005; Dayan, 2008). Cooperation among, and competition between, the different systems influence the responses of subjects in experiments, although the exact interactions are only beginning to become clear.

In this chapter, we first outline the components of this architecture, focusing on different systems involved in evaluating outcomes and choosing actions. Their impact in the broader field of neuroeconomics has historically been most apparent in positive cases; we therefore focus on two key asymmetries between loss-related and reward-related issues. One of these relates directly to the anomalies of choice listed above, and arises from the influence on normative, reward-maximizing, and punishment-minimizing choices of innate responses to aversive predictions and outcomes. That the mere prediction of an aversive outcome can have an effect on behavior that, paradoxically, increases the chance of attaining that outcome is an Achilles heel of decision making with widespread unfortunate consequences (Breland and Breland, 1961; Dayan *et al*., 2006).

The second asymmetry has to do with learning. In one important class of tasks, subjects are penalized for any action they take except for one, particular, choice (selected ahead of time by the experimenter). For such tasks, telling subjects that they just performed a bad action does not, in general, tell them what they could have done instead that would have been better. By contrast, telling them that an action was not bad is much more specifically useful. The consequence of this asymmetry lies in psychological and neural aspects of the interaction between learning associated with rewards and punishments. Learning which actions to execute to avoid punishments appears to require the involvement of positive signals created through mutually opponent interactions between separate systems involved in appetitive and aversive predictions and outcomes. The positive signal arises in the light of the progression from a state in which punishment is expected to a state in which it is not. Although this asymmetry has fewer direct consequences for existing economic tasks, for which learning often plays a somewhat restricted role, it is important in ecologically more natural settings.

We start by describing the architecture of prediction and decision making in positive and negative contexts. We then discuss a class of so-called Pavlovian influences over choice in negative contexts; and finally consider issues to do with learning. Loss aversion itself is discussed in detail elsewhere (see Chapter 11).

THE ARCHITECTURE OF AFFECTIVE DECISION MAKING

The fields of economics, operations research, control theory, and even ethology share a common theoretical framework for modeling how systems of any sort can learn about the environments they inhabit, and also can come to make decisions that maximize beneficial outcomes and minimize adverse ones (Mangel and Clark, 1988; Puterman, 1994; Camerer, 1995; Sutton and Barto, 1998). This framework is closely associated with dynamic programming (Bertsekas, 1995) and encompasses many different algorithmic approaches for acquiring information about an unknown environment, including learning from trial and error, and using that information to specify controls. It has recently become apparent that different structures in the brain instantiate various of these approaches, in some cases in a surprisingly direct manner; producing a complex, but highly adapted and adaptive overall controller (Dickinson and Balleine, 2002; Daw *et al*., 2005; Dayan, 2008).

In many cases for experimental and behavioral economics, the specification of the problem includes exactly the full costs and benefits of each course of action in a stylized tableau (Camerer, 1995). However, in typical natural cases of decision making, this simplifies away at least two issues. First, feedback for a choice is usually only available after some time has elapsed and, potentially, also additional choices (as, for instance, in a maze). This problem of delayed feedback seems to have played an important role in determining the nature of the neural controllers, with forms of prediction lying at their heart (Montague *et al*., 1996; Sutton and Barto, 1998). The second main difference between natural and economic decision making is that the latter mostly involves money, which only has derived, and not intrinsic, value to the subjects. The extent to which proxies such as money, let alone more abstract outcomes such as mere points in a computer game, can entrain neural decision-making structures

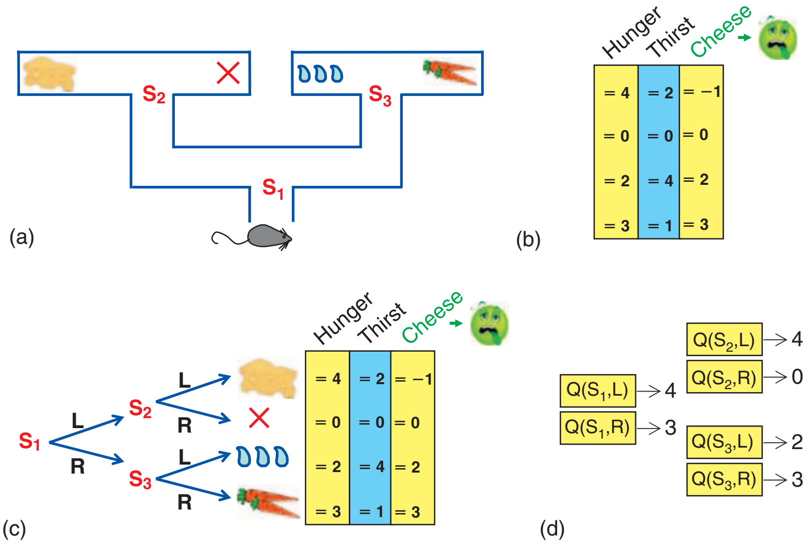


FIGURE 12.1 Model-based and model-free actions in a simplified maze task. (a) A simple maze with three states (S1, S2, and S3) from which the animal has to make left–right decisions, with the terminal states yielding outcomes of cheese, nothing, water or carrots. (b) The values of these outcomes under three different motivational states: hunger, thirst, and cheese devaluation. This latter state results from cheese ingestion with vomiting (artificially induced by lithium chloride injection in most experiments). (c) A tree-based model of the state-action environment, which can be used to guide decisions at each state by a model-based controller. (d) The cached values available to a model-free, habitual controller. Immediately after cheese devaluation, these values do not change (in contrast to the model-based controller). It is only after direct experience with the devalued cheese that the value associated with Left (S2), and subsequently Left (S1), is reduced. Figure adapted from Niv *et al*. (2006).

that are presumably evolved to handle natural rewards (“reinforcers”) such as food, water, and intrinsic threats is actually quite remarkable.

The essence of the solution to the problem of delayed feedback is prediction of the value of being in a particular situation (typically called a *state*) and/or doing a particular action at that state, in terms of the rewards and punishments that can be expected to accrue in the future. Different ways of making predictions underlie different approaches to control, leading to an overall architecture that is complicated. In particular, we have suggested that there is evidence for at least four different sorts of predictor or value system and four different sorts of controller (Dayan, 2008). However, for the present purposes, two predictors and three associated controllers are most important.

The predictors (called *model-based* and *model-free*, for reasons that we discuss below) trade off the complexity of learning for the complexity of computation. These predictors are directly associated with two of the controllers (which psychologists refer to respectively as goal-directed and habitual). The third controller (called *Pavlovian*) uses the model-based and model-free values, but emits responses that are selected by evolution rather than learning. We argue that the Pavlovian controller plays a critical role in creating decision-theoretic anomalies (Dayan *et al*., 2006).

In the rest of this section, we describe these key value systems and controllers. We organize the descriptions around the simple rodent maze task shown in Figure 12.1a (adapted from Niv *et al*., 2006). This has three choice points (A, B, and C); and four possible outcomes (cheese, nothing, water, and carrots). When the animal is hungry, the cheese is most valuable – i.e., has the highest outcome utility – followed by the water and carrots; when thirsty, the water is most valuable. However, the cheese can be devalued, either by allowing the animal to eat it freely until it chooses to eat it no more (this is called *sensory-specific satiety*, since the value of the cheese is specifically lowered) or by injecting the animal with a chemical (lithium chloride) after it eats some cheese. The latter treatment makes the animal sick, an outcome that induces a form of specific food aversion, such that, again, the cheese is no longer valuable. Figure 12.1b shows the utilities of each of the outcomes under the three motivational states of hunger, thirst, and cheese aversion.

Model-based Values; Goal-directed Control

One obvious way for a subject to make predictions about future punishments or rewards is to use a model of the world. This model should indicate the probability with which the subject will progress from one state to the next, perhaps dependent on what actions it takes, and what the likely outcomes are at those states, which again may depend on the actions (Sutton and Barto, 1998). Figure 12.1c depicts the model of the simple maze task; it is nothing more than the tree of locations in the maze, which are the states of the world, joined up according to the actions that lead between them. Not only does the model specify which outcomes arise for which actions; it should also specify the (expected, experienced) utility of those outcomes. As shown in the figure, this depends on the motivational state of the subject. The information necessary for the model can readily be acquired from direct experience in an environment, at least provided that the environment changes at most relatively slowly.

Given some systematic way of acting at each location or state (e.g., choosing to go left or right with probability 0.5), models such as that shown in Figure 12.1c admit a conceptually very simple way of making predictions about the values of states or locations in the maze, namely searching forward in the model, accumulating expected values all the while. Unfortunately, computing model-based values accurately when there are many different possible states and actions places a huge burden on working memory, and also on aspects of calculation. The values can therefore only possibly be accurate in rather small environments.

Of course, it is not enough to compute the value of a random choice of action at a location; rather, it is necessary to find the best action. Since the model in Figure 12.1c actually specifies the utility consequences of the different possible actions, it can also straightforwardly be used to perform the dynamic programming step of finding the optimal action. This can, in principle, be performed either forwards or backwards in the tree.

One critical facet of this model-based method of choosing actions is that the decision utilities used to make choices, i.e., the information about the expected utilities of the actions, can depend on a calculation as to which outcomes will result, and what their expected (experienced) utility will be. Take the case that the model includes all the utilities shown in Figure 12.1b. If the subject is trained whilst hungry, it will normally turn left at A to get the cheese. However, as soon as the cheese has been devalued through pairing with illness, the prediction of the utility of going left at A will be reduced, and the subject will turn right instead to get the carrots.

In psychological terms, since these values depend on the expected outcomes and their modeled utilities, this sort of control is considered to be goal-directed (Dickinson and Balleine, 2002) since these utilities define the animals' goals. This sort of outcome-sensitive control is a close relative of human notions of "cognitive" control, in which individuals explicitly consider the outcome of actions and of subsequent actions, and use some form of tree-search to inform current actions. The brain might support different ways of doing this – for instance, using propositional, linguistic structures or, by more or less direct analogy with navigation, structures associated with spatial processing. It is closely related to the classical notion of outcome-expectancy expounded by Tolman (Tolman, 1932). Indeed, model-based prediction and control has the key characteristic of being highly flexible over the course of learning – new information about the environment can be fit into the model in exactly the right place to have an appropriate effect.

Further, it might be imagined that subjects could acquire higher-order information about the overall structure of the environments they experience that might generalize from one particular task to another. One example that has been highly influential in the psychological literature is that of controllability – a measure of the influence a subject might expect to have over its outcomes. There is a range of experiments into what is known as *learned helplessness* (Maier and Seligman, 1976) in which subjects are taught that they cannot control some particular aspect of one environment (for instance, being unable to influence a shock). They generalize this inference to other environments, failing to explore or exploit them effectively. There are various possible formalizations of controllability as Bayesian prior distributions over characteristics of the models (Huys and Dayan, 2008), but more data are necessary to pin this issue down completely.

The neural instantiation of the model and associated calculations for predictions and action choice is not completely known. However, there is evidence for the involvement of several regions of prefrontal cortex, including ventromedial prefrontal cortex (related to the prelimbic and infralimbic cortex in rats), lateral orbitofrontal cortex, and middle frontal gyrus, along with the dorsomedial striatum (Balleine and Dickinson, 1998; Dayan and Balleine, 2002; Koechlin *et al.*, 2003; Ursu and Carter, 2005; Carter *et al.*, 2006; Yin *et al.*, 2006; Yoshida and Ishii, 2006). Most of these experiments involve rewards rather than punishments, though, and the representation of model-based negative values is not wholly clear.

Model-free or Cached Values; Habitual Control

The problem with model-based prediction and control is the complex, and thus error-prone, calculations that are necessary to compute values. One way round at least some of this complexity is to collapse the total anticipated value of future state transitions or actions by storing (or, to use a word taken from computer science, caching) what would be the results of this tree search (Sutton and Barto, 1998; Daw *et al*., 2005). In effect, a cached value provides a single simple metric, an outcome-independent neural currency, as to the overall utility of a particular state, or taking a certain action at that state.

Figure 12.1d shows the cached values (called *Q-values*) of each action at each location in the maze, assuming that the subject chooses optimally for the state of hunger. Such cached values can be used without direct reference to a model of transitions or outcomes; hence this form of prediction is often termed model-free. These values are represented by a function (the Q function) whose argument is the state (here, the location in the maze).

Of course, the cached values in Figure 12.1d are just the same as the optimal values produced by model-based evaluation in the case of hunger. However, critically, it turns out that these values can be learned directly over the course of experience of state transitions and utilities, without any reference to a model at all. Ways to do this, i.e., ways of implementing asynchronous, sampled, dynamic programming, are highlighted in Chapters 3, 22 , 23, 24, and 26 of this volume under the guise of temporal difference methods of reinforcement learning (Sutton and Barto, 1981; Barto *et al*., 1990; Watkins and Dayan, 1992). Temporal difference learning works by exploiting the key property possessed by the cached values in Figure 12.1d – namely, consistency from one state to the next. For example, since no outcome is provided at state A, the value of going left at that state is just the same as the value of the state (B) consequent on going left there; the value of going right is the same as the value of the state (C) that arises for going right. The discrepancy (if any) between these successive value estimates is exactly the basis of the temporal difference learning rule. In this way, sequential estimates of values effectively transfer between adjacent states, obviating the need to wait for actual outcomes themselves.

Perhaps surprisingly, it turns out that temporal difference algorithms are not just distant abstractions over baffling neural complexities. Rather, at least in the case of positive outcomes, there is substantial evidence (also reviewed in Chapters 3, 21, 24, and 26 of this volume) that the moment-by-moment (phasic) activity of cells that project the neuromodulator dopamine to the striatum matches closely the key prediction error term in temporal difference learning, providing a signal that is ideally suited for manipulating predictions appropriately (Montague *et al*., 1996; Schultz *et al*., 1997; Satoh *et al*., 2003; Nakahara *et al*., 2004). Unfortunately, the case of aversive outcomes is less well understood. fMRI studies suggest that punishments lead to prediction errors with rather similar properties to those for rewards (Seymour *et al*., 2004; Jensen *et al*., 2007), although electrophysiological evidence from animals is thinner on the ground (Belova *et al*., 2007).

Most importantly for model-free predictions is that the brain appears not to use the obvious representation in which rewards (and positive prediction errors) are coded by greater-than-average neural activity in a neural population, and losses (and negative prediction errors) by less-than-average neural activity in the same population. Rather, as in many other cases, it seems to use two systems that oppose each other (Konorski, 1967; Solomon and Corbit, 1974; Dickinson and Dearing, 1979; Grossberg, 1984; Seymour *et al*., 2005, 2007a). In this arrangement, positive outcomes can inspire responses from the negative system when they are unexpectedly omitted, or when sequences of them cease. Further, stimuli which predict the absence of rewards (called *appetitive inhibitors*) and stimuli which predict the presence of punishments or loss (aversive excitors) are treated in a formally similar manner. For example, in terms of value representations, omission of food is intrinsically similar to painful shocks. This is demonstrable in various psychological paradigms (Dickinson and Dearing, 1979). Conversely, there is a natural similarity between appetitive excitors and aversive inhibitors.

The neural realization of the system associated with negative, model-free values that is opponent to dopamine is not completely resolved. One class of theoretical models hints at the involvement of a different neuromodulator, 5-hydroxytryptamine (5-HT or serotonin), as a more or less direct opponent (Daw *et al*., 2002). However, direct evidence for this possibility is scant, there are competing theories for the role of this neuromodulator, and the fMRI studies, with their poor spatial resolution and the uncertainties about exactly what aspects of neural activity they capture in structures such as the striatum (Jensen *et al*., 2003, 2007; Seymour *et al*., 2005, 2007a), leave us without a completely unified picture.

In fact, until recently the striatum had been considered to be reward-specific in economic studies in humans. However, the findings above, and others (Seymour *et al*., 2004, Delgado and colleagues, forthcoming), along with ample animal studies (Ikemoto and Panksepp, 1999; Horvitz, 2000; Schoenbaum and Setlow, 2003; Setlow *et al*., 2003; Wilson and Bowman, 2005) suggest that the striatum is involved in both appetitive and aversive processing, and indeed may be a critical point in the brain where these opposing motivational streams are integrated. Slightly clearer is the representation of the cached aversive values themselves, which evidently involves the amygdala and anterior insula cortex (Seymour *et al.*, 2004; Paton *et al.*, 2006).

The clear advantage that model-free, cached values have over model-based values is that they are represented directly, and do not need to be computed by a process of tree-based evaluation that imposes a heavy burden on working memory, and is likely to be inaccurate in even moderately complex domains. However, attending this computational benefit is statistical inefficiency over learning, and inflexibility in the face of change.

First, the drive underlying temporal difference learning is discrepancy between the predictions made at successive states. However, early in learning, the predictions at all states are wildly inaccurate, and therefore the discrepancies (and thus the temporal difference prediction error) are of little use. Thus, model-free learning is statistically inefficient in the way it employs experience. To put the point another way, temporal difference learning involves bootstrapping (i.e., using one estimate to improve another one), a procedure which is far from optimal in its use of samples from the environment.

The second problem with model-free methods is inflexibility. As we noted, cached values such as those shown in Figure 12.1d are just numbers, divorced from the outcomes that underlie them, or the statistics of the transitions in the environment. This is why caching is computationally efficient. However, if the motivational state of the subject changes (for instance if the cheese is poisoned, as in the rightmost column of Figure 12.1b), then the cached values will not change without further, statistically expensive, learning. By contrast, the model-based values, which are based on direct evaluation in the tree of outcomes, can change directly.

In Figure 12.1d, the model-free values are predictions of the long-run utilities of particular actions at each location. They can thus be directly used as decision utilities, to choose between the possible actions at a location. This leads to a model-free controller, one that makes decisions without reference to a model of the environment. We pointed out above that the cached values do not change with the motivational state of the subjects without further learning, and so the model-free decisions will not change either. In psychological terms this is exactly the characteristic of habits (Dickinson and Balleine, 2002), and so this controller is deemed habitual (compared with the goal-directed control associated with the model-based value system).

From a neural perspective, there is evidence for the involvement of the dorsolateral striatum in representing the values of taking actions at states (Yin *et al.*, 2006), and indeed in habitual control. In the appetitive case, again, dopaminergic projections from the substantia nigra pars compacta to this region of the striatum are believed to play a central role in learning (Montague *et al.*, 1996; Schultz *et al.*, 1997). The habits themselves may be represented or stored in cortico-thalamic loops (Yin and Knowlton, 2006).

The habitual controller defined above involves the competition between different actions depending on values (or other quantities depending on the values) that are the output of a function of the state (the Q function in Figure 12.1d). An even more primitive form of habitual controller would use a function to parameterize the mapping from state to action directly, without going through the intermediate value of a range of actions (Barto *et al*, 1983). Psychologists would consider this to be a stimulus (i.e., state)-response (i.e., action) mapping. It is also model-free, and insensitive to motivational changes, and thus hard to distinguish behaviorally from the Q-value-dependent, model-free controller described above. There are intriguing reports of just such a controller in even more dorsolateral striatal regions (Everitt and Robbins, 2005).

The existence of multiple controllers (goal-directed and habitual) gives rise to a new choice problem – that of choosing between them. One view is that they compete for the control of behavior according to their relative uncertainties (Daw *et al.*, 2005). Model-based values are favored early in the course of learning, because of their greater statistical efficiency. However, model-based values are disdained once sufficient samples have accumulated, because the computational demands of calculating them inevitably lead to extra noise.

Most work distinguishing habitual and goal-directed control has involved appetitive outcomes, and we discuss some subtleties of aversive habitual control later in the chapter.

Pavlovian Control

Model-based and model-free controllers can, in principle, learn arbitrary actions to optimize their behavior, at least those actions that can be expressed and explored. Indeed, these are often referred to as instrumental controllers, since their choices are learned to be instrumental for the delivery of desired outcomes. Although this flexibility is very powerful, it comes with an attendant cost of learning. Evolution appears to have endowed everything from the simplest organisms to us with powerful, pre-specified, but inflexible alternatives (Konorski, 1967; Dickinson, 1980; Mackintosh, 1983). These responses are termed *Pavlovian*, after the famous Russian physiologist and psychologist Pavlov.

Immediately available rewards, such as food or water, and immediate threats, such as pain or predators (collectively called *unconditioned stimuli*), elicit a range of apparently unlearnt, typically-appropriate, so-called consummatory, responses. For appetitive outcomes these are relatively simple, although they may reflect certain specific attributes of the outcome – for instance, differing for solid and liquid outcomes. The consummatory responses associated with aversive outcomes appears to be more sophisticated than those for rewards, including increased heart rate and sweating during acute pain, fighting in the midst of a contest, and leg flexion in the face of foot-shock. The choice between the whole range of defensive and aggressive responses depends rather precisely on the nature of the outcome, the context, and particularly the effective ("defensive") distance of the threat (Blanchard and Blanchard, 1990). These responses are seemingly under the control of a brainstem structure, the periaqueductal gray (PAG), which has a rich, topographically organized architecture (Fanselow, 1994; Fendt and Fanselow, 1999; Graeff, 2004; Mobbs *et al*., 2007).

However, and more relevantly for us, predictions associated with these appetitive or aversive outcomes also elicit an often somewhat different set of so-called *preparatory responses*. These are automatically tied to the predictions, independent of whether they are actually appropriate responses in a given circumstance. They thus provide an additional route by which the predictive mechanisms discussed in the previous subsections can generate behavior.

Such preparatory responses are also varied. For instance, in rats, anticipation of a shock causes attempted escape if the cue underlying the anticipation is localized at a particular point in the environment (e.g., a light LED), but freezing if it is more general. Such anticipation can also lead to fighting in the presence of another male, and copulation in the presence of a female (Ulrich and Azrin, 1962; Sachs and Barfield, 1974). However, there are also preparatory responses that reflect the general positive or negative valence of the predicted outcome, and elicit non-specific responses such as approach or withdrawal. We suggest in the next section that it is these general preparatory responses, arising largely from predictions of financial gain and loss, that are associated with significant behavioral anomalies in human economic choices.

The neural realization of Pavlovian responses has been well studied (LeDoux, 2000; Maren and Quirk, 2004). As mentioned above, aversive value predictions depend critically on the amygdala (Cardinal *et al*., 2002; Balleine and Killcross, 2006). The amygdala is a complex and incompletely understood structure with many sub-parts. However, it seems that one subarea, the central nucleus, is predominantly involved in directing non-specific preparatory responses. These include arousal and autonomic responses, and also approach/withdrawal, achieved through its extensive connections to brainstem nuclei and one part of the nucleus accumbens (termed the *core*). Another sub-area, the basolateral complex, is predominantly involved in much more specific responses, mediated downstream through connections to regions such as the hypothalamus and periaqueductal gray and a separate part of the nucleus accumbens (termed the *shell*).

PAVLOVIAN INFLUENCES OVER INSTRUMENTAL BEHAVIOR

The responses of the Pavlovian controller are determined by evolutionary (phylogenetic) considerations rather than (ontogenetic) aspects of the contingent development or learning of an individual. These responses directly interact with instrumental choices arising from goal-directed and habitual controllers. This interaction has been studied in a wealth of animal paradigms, and can be helpful, neutral or harmful, according to circumstance. Although there has been less careful or analytical study of it in humans, we have argued that it can be interpreted as underpinning a wealth of behavioral aberrations (Dayan *et al*., 2006).

Crudely, predictions of future appetitive outcomes lead to engagement and approach; predictions of future aversive outcomes lead to disengagement and withdrawal. For instance, consider the phenomenon of Pavlovian-instrumental transfer (PIT) (Estes, 1948; Lovibond, 1983; Dickinson and Balleine, 2002). In this, the speed, rate, alacrity or, more generally, vigor with which subjects perform an instrumental response for a particular positive outcome is influenced by the mere presentation of stimuli that are associated in a Pavlovian manner with either appetitive or aversive outcomes. In our terms, the stimuli signal states, the important aspect of PIT is that the predictive association of the Pavlovian stimulus occurs separately from that of the instrumental context.

Vigor is boosted the most by stimuli predicting an appetitive Pavlovian outcome that is exactly the same as the outcome of the instrumental behavior. This so-called specific PIT depends (at least in rats) on the integrity of the basolateral amygdala and nucleus accumbens shell (Corbit *et al.*, 2001; Cardinal *et al.*, 2002; Corbit and Balleine, 2005). However, vigor is also boosted by stimuli predicting motivationally relevant appetitive outcomes (such as water, for a thirsty subject) that are different from the instrumental outcome. This is called *general PIT*, and may be seen as a general, non-selective, preparatory appetitive phenomenon. In rats, general PIT depends on the integrity of the central amygdala and nucleus accumbens core (Corbit *et al.*, 2001; Cardinal *et al.*, 2002; Corbit and Balleine, 2005), in keeping with the description above about the neural realization of Pavlovian conditioning.

Finally, stimuli predicting aversive Pavlovian outcomes can actually suppress appetitive instrumental responding, and lead to extraneous actions such as withdrawal. This is normally called *conditioned suppression* (Estes and Skinner, 1941) rather than aversive PIT, which would perhaps be the more natural term. However, it is a ubiquitous and powerful phenomenon that is in fact often used as a sensitive measure of the strengths of aversive Pavlovian predictors.

Most critically, choice, as well as vigor, is affected by these Pavlovian predictions. This is seen very clearly in a slightly complex paradigm called *negative automaintenance* (Williams and Williams, 1969). In one example of negative automaintenance, pigeons are shown the predictive association between the illumination of a key and the delivery of food. The Pavlovian prediction associated with the lighting of the key automatically elicits a peck response on the key, as a form of preparatory approach and engagement. In fact, this part of the procedure is one of the standard forms of Pavlovian conditioning, which is termed *autoshaping* because of the automaticity of the pecking (Brown and Jenkins, 1968). The experimenter then arranges what is called an *omission schedule*, so termed because on any trial in which the pigeon pecks the key when illuminated, no food will be provided. In this case, there is a battle between the Pavlovian response of pecking and the instrumental need to withhold. Pigeons cannot help themselves but peck to some degree, showing the critical, and indeed in this case, deleterious, impact of the Pavlovian prediction.

Although it has been suggested that Pavlovian responses interfere comparatively more with instrumental habits than goal-directed actions, the factorial PIT-based interactions between model-based and model-free Pavlovian predictions and model-based and model-free instrumental actions have not been systematically studied.

There appear to be fewer aversive examples of phenomena like negative automaintenance, which is somewhat surprising given the robustness of Pavlovian aversive responses in general. Where they can be shown to exist, they yield self-punitive behavior. In one putative example, squirrel monkeys were punished, by way of an electric shock, for pulling on a restraining leash (Morse *et al.*, 1967). The (instrumentally) optimal action in such a circumstance is to stop pulling; however, one Pavlovian response to shock in the time leading up to its expected delivery is to try and escape by pulling. As expected from Pavlovian misbehavior, the monkeys did in fact pull on the leash more rather than less. A similar example is seen in Siamese fighting fish, who can be trained to swim through a hoop and perform an aggressive display. If an experimenter then tries to inhibit this display by an aversive shock, the behavior is paradoxically augmented (Melvin and Anson, 1969). This is most likely since the aggressive display is part of the innate repertoire of defensive responses, which turns out to be extremely difficult to overcome.

What, then, are the neuroeconomic consequences of these Pavlovian effects? After a methodological note, we briefly consider four: impulsivity, framing, depressive realism, and dread. Note that these are all complex and rich phenomena; we only focus on the subset of issues that Pavlovian control may explain. This may seem like the same sort of smorgasboard of issues to which other broad explanatory frameworks such as hyperbolic discounting have been turned; rather, we argue that it is critical to understand the breadth of phenomena associated with something as basic as Pavlovian conditioning, given its overwhelming evidentiary basis in psychology and neuroscience.

Methodology

We must first raise a couple of methodological points about the relationship between economic and psychological paradigms. In experimental and behavioral economics, decisions are often probed in relation to options with stated parameters – that is, the magnitudes, risks and uncertainties of various options are given directly. These are likely to exert their effects mostly through model-based predictions (and goal-directed control). By contrast, in experimental psychology, the parameters of options are typically learned through trial and error. Thus, representations of value and risk are experience-based rather than propositional, and can have an impact through model-free as well as model-based control. Of course, experience-based representations are imperative in animal experiments, and have also been highly successful in deconstructing the components of aversive (and appetitive) behavior. However, any complete account of aversive behavior needs to integrate both, since humans are presented with both types of situation: one shot decisions, such as those regarding pensions and life insurance; and repeated decisions, such as those regarding what painkiller to take or which foods to buy.

A further difference in methodologies relates to type of aversive events used. Neuroscientists have often used pain, for instance in the form of an electric shock to hand or paw. The advantage of this is it is an immediately and relatively instantaneously consumed commodity. Furthermore, it is both potent and ecologically valid, in the sense that it is the sort of stimulus with which aversive systems evolved to deal. We should therefore say a word about the neural processing of pain itself.

Physical pain is subserved by a sophisticated system of specialized neural pathways signaling information about actual or imminent tissue damage to many areas of the spinal cord and brain (Julius and Basbaum, 2001; Craig, 2002; Fields, 2004). This results not just in the set of characteristic, involuntary, defensive responses described above, but also in a perceptual representation of negative hedonic quality.

In the brain, the basic representation of aversive innate value implicates brainstem and midbrain structures, including the periaqueductal gray, parabrachial nucleus, and thalamus (Lumb, 2002). Cortical structures such as insula (particularly anterior regions), lateral orbitofrontal and mid-anterior insula cortices are more directly associated with refined aversive representations, including conscious negative hedonic experience (Craig, 2002). These correlate more closely with the subjective experience of unpleasantness, which in humans often accompanies innate aversive outcomes. In fact, the feeling associated with loss dictates the way these systems are often described in traditional psychological accounts (Price, 1999). This can, however, be approached more formally by considering "feeling" as a process of hedonic inference. As with many less motivationally-laden sensory systems, afferent information is rarely perfect, and a statistically informed approach is to integrate afferent input with either concomitant information from other modalities (multi-sensory integration), or prior knowledge of events (expectation).

By contrast with these rich phenomena associated with actual threats, economists have, naturally, tended to use financial losses. Various of the other chapters in this volume capture aspects of the psychological and neural richness of money as a stimulus; for simplicity, we adopt the straightforward view of it as a conditioned reinforcer, that is, a stimulus that has undergone (extremely extensive) Pavlovian training to be associated with many different sorts of future reward. In these terms, losing money is like taking away a conditioned reinforcer; an outcome that is indeed known to be aversive.

One complicating issue is the slightly unclear relationship between the affective values of states and those associated with state changes (Kahneman and Tversky, 2000). Take a concrete example – the state of hunger. On the one hand, this would seem to be clearly an aversive state – it poses a threat to homeostasis. On the other, the affective worth of the same morsel of food is greater when hungry than when sated, and so, for instance, the average long-run experienced utility may actually be higher (Niv *et al.*, 2006). Is the apparently masochistic act of starving yourself actually utility maximizing in that you enjoy food in the future sufficiently more? In general, teasing apart the contribution to utility of the actual outcome and the motivational state within which it is evaluated is hard.

The answer to the masochism question is not yet quite clear. However, it does pertain to one of the other value systems that we have not yet discussed. Most economic decision-making tasks are one-shot or phasic. By comparison, many psychological paradigms for animals are ongoing or continuous. For these, it often makes sense to predict and maximize the long-run average rate of rewards rather than, for instance, the more conventional long-run sum of exponentially-discounted rewards. In this case, this average rate of reward has a status as something like an opportunity cost for time. Niv and colleagues (2007) noted this, and studied a framework in which subjects were free to choose not only which actions to do, but also how fast to do them. Under the reasonable assumption that acting quickly is expensive, it turns out that the optimal speed or vigor of responding is determined by the average rate of reward. Arguing partly on the basis of the data on the control of vigor from the Pavlovian-instrumental transfer paradigms we discussed above, they suggested that the long-run, tonic, level of dopamine or dopaminergic activity should report this average reward. This is the additional value system. However, vigor is also important in cases in which signaled punishments or aversion can be avoided through active actions. Tonic dopamine may therefore represent the sum of average rewards and avoidable punishments; bar the expectation of a long-run absence of food, hunger is exactly an example of this sort of case. Whether the tonic aversiveness of hunger is also represented by the tonic activity of another system (for instance, some subset of 5-HT cells) is not clear.

For the present, we will just consider phasic aversive outcomes, such as shocks, or immediate financial losses, together with predictions of these. Neurobiological evidence is starting to accrue that confirm that the underlying motivational processes in financial loss share strong similarities with that associated with physical pain (Delgado *et al.*, 2006; Knutson *et al.*, 2007; Seymour *et al.*, 2007a). For example, Knutson and colleagues have suggested that financial amounts associated with payments in shopping transactions are correlated with activity in and around insula cortex (Knutson *et al.*, 2007), which has also been shown to correlate with expected value of pain (Seymour *et al.*, 2004). We have shown activation to prediction errors for financial loss in striatum, in a similar manner to those seen in studies of aversive conditioning for painful shocks (Seymour *et al.*, 2007a). Delgado and colleagues (forthcoming) have recently shown directly the common striatal aversive processing for pain and financial loss, by engaging subjects in a task that involves both.

Impulsivity and Altruistic Punishment

Impulsivity covers a broad range of phenomena. Classically, it features engagement in actions whose immediate benefits are less than those of longer term payoffs that would accrue if the subjects could be patient (Cardinal *et al.*, 2004). That is, subjects exhibit temporal short-sightedness. Impulsivity is best described in the appetitive domain, but similar notions may apply in aversive domains too. In the appetitive case, we have argued that the effect of a Pavlovian approach response associated with a proximally available beneficial outcome can be to boost early, and thus impulsive, responding at the expense of what would be favored by goal-directed or habitual instrumental systems (Dayan *et al.*, 2006). Treating this form of impulsivity in Pavlovian terms amounts to a subtly different explanation of the behavior from accounts appealing to (or data fitting with) hyperbolic discounting or indeed ideas about differences between (model-based) rational and (model-free or perhaps neuromodulator-based) emotional cognition, which conventionally ignore the normative intent of model-free control.

In the aversive case, one example of apparent impulsiveness is altruistic punishment, in which subjects punish others (typically free-riders who fail to cooperate in various forms of group interactions, but nevertheless take advantage of the group effort) at a pure cost to themselves (i.e., with negative immediate benefit), without any prospect of a direct return on this investment of effort or risk (i.e., with no long term payoff at all). Although the nature of the actions which subserve altruistic punishment remain unclear (Seymour *et al.*, 2007b), there is good evidence that humans readily engage in such actions (Fehr and Gachter, 2002; Yamagishi, 1986; see also Chapter 15 in this volume).

Certainly, some aspects of apparent altruism can be explained by reputation formation (a form of indirect reciprocity) and tit-for-tat (a form of direct reciprocity). These can be captured by model-based and even model-free instrumental mechanisms. The argument that altruistic punishment is partly a Pavlovian anomaly is that (a) punishment is a form of aggression, whose innate roots we explored above, and (b) in highly social species such as humans, there is an evolutionary imperative to prevent exploitation by free-riders that is satisfied by making non-cooperation expensive. First, innate aggression is evidently a potentially life-saving mechanism of defense in the face of predators, and in within-species contests, can be important for protecting food, territory and mating partners (Clutton-Brock and Parker, 1995). Second, in humans, and possibly some other primate species, aggressive responses can also serve to promote cooperation, since they provide a negative incentive for members of a group to exploit each other, and protect various forms of reciprocity (Boyd and Richerson, 1992; De Waal, 1998; Stevens, 2004). Thus innate responses to perceived unfairness may have evolved on the basis of punishment in these sorts of non-altruistic circumstances, such as in groups or societies of small enough size such that individuals (and certainly their kin) would be likely to interact repeatedly with offenders, rendering the punishment non-altruistic (i.e. "selfish"). However, once established as an innate response, punishing non-cooperators could have become blind to its proximal consequences for the individual (like other Pavlovian responses), thus appearing impulsive.

There is also the alternative possibility that altruistic punishment arises from the structural inefficiency of instrumental control associated with habits, rather than the interference of Pavlovian imperatives over instrumental ones. Crudely, the idea is that choosing precisely who to punish in a circumstance requires the detailed calculations of the consequences of punishment and likelihood of future interactions that only the goal-directed system could entertain. However, the habit system can engage in instrumental punishment in reciprocal cases and may therefore gain control over all similar such conditions, as discussed above. Its inability to calculate in detail the consequences of its output can then lead it to punish "inappropriately" in altruistic situations. This type of "error" resembles that seen in devaluation experiments, when habitually trained animals fail to reduce responding to outcomes that have been separately paired with punishment.

Framing Effects

Framing effects are a rather well-studied peculiarity of human (and non-human; see Chapter 7 of this volume) choice in which the decision between options is influenced by subtle features of the way in which those options are presented. Typically, the language used to describe an option is manipulated in a valance related manner, whilst the expected value remains unchanged.

The so-called "Disease dilemma" is a popular example. In this, subjects are asked to choose between two options relating to the management plan of an epidemic, one of which contains risk and the other not (Tversky and Kahneman, 1981). The risky option is fixed, such as "Option A has 2/3 chance of curing all 600 affected people", but the non-risky option is presented in either a positive or negative frame, as either "With Option B, 400 people will be *saved*" or "With Option B, 200 people will *die*". Subjects tend to choose the risky option when the sure option is presented in terms of people dying, and the sure option when presented in terms of the numbers who will be saved.

Similarly, De Martino *et al*. (2006) conducted a study involving loss/gain framing of non-risky, alongside risky, financial options, matched for expected value. Subjects showed a risk preference reversal from risk aversion to risk-seeking when the choice was switched to a loss frame. This change in behavior was positively correlated with amygdala activity.

Given the role of the amygdala in Pavlovian-instrumental transfer, and thus the untoward influence of predictions on instrumental actions (Corbit and Balleine, 2005), results such as this are consistent with a Pavlovian component to framing. That is, an option which is presented as involving sure deaths will automatically engage a Pavlovian aversive withdrawal response decreasing its propensity to be chosen, that is absent for the option involving sure survival. The latter might generate an appetitive approach response instead. As we have seen above, model-based evaluation mechanisms, which could compute the equality between the options, are not the only source of predictions; model-free mechanisms, which lack such computational power, also exert their influence, in this case in just the direction shown. Indeed, we can look at the classic trolley moral dilemmas (Thomson, 1986) in a similar light. Even if subjects didn't have any choice, but just had to execute an action to register a single option, we would predict that the same Pavlovian effect would make their reaction times slower, an effect seen in other experiments (Shidara *et al*., 2005; Sugase-Miyamoto and Richmond, 2005).

Depressive Realism

In comparisons between healthy volunteers and patients with depression, a (not completely uncontroversial) finding is that the volunteers are unduly optimistic about the appetitive value of, and the degree of control they exert over, artificial, experimentally-created environments. By contrast, the depressed subjects make more accurate assessments, and so are more realistic. This phenomenon is called *depressive realism* (Abramson *et al*., 1979). Further, by comparison with control subjects, depressed patients ruminate on negative outcomes.

It has been suggested that Pavlovian withdrawal associated with predictions of negative outcomes is an important route to the over-optimism of the volunteers, and that one of the underlying neural malfunctions associated with depression is associated with a weakening of this withdrawal, thereby leading to more accurate, but more pessimistic, evaluations (Huys and Dayan, 2008). Consider a healthy subject entertaining chains of thought about the future. Any chain of thought leading towards a negative outcome engenders a Pavlovian withdrawal response, which may lead to its being terminated or (in the jargon of tree-based search) pruned. Thus, if healthy subjects contemplate the future they will tend to favor samples with more positive outcomes, and will therefore be more optimistic. Given the possibility that this form of Pavlovian withdrawal is mediated by 5-HT, as the putative aversive opponent to dopamine (Daw *et al*., 2002), and the pharmacological suggestion that depressed patients have low effective 5-HT levels (Graeff *et al*., 1996), it is conceivable that this withdrawal mechanism is impaired in the depressed subjects. This would, of course, lead to the basic phenomenon of depressive realism. Indeed, boosting 5-HT, which is the ultimate effect of the standard treatment for depression, namely selective 5-HT reuptake inhibitors, helps restore the original optimism.

Altered levels of 5-HT are also associated with other phenomena, such as impulsivity (Cardinal, 2006; Chamberlain and Sahakian, 2007), which have been argued to have Pavlovian roots.

Dread

In an aversive domain, many subjects show an additional sort of impulsivity in the form of dread (Berns *et al.*, 2006). They prefer a larger electric shock that comes sooner to a weaker shock that comes later, reportedly because of the misery of aversive anticipation (Loewenstein, 1987, 2006; Caplin and Leahy, 2001). Indeed, during the anticipation phase in the study by Berns and colleagues, brain regions commonly associated with physical pain are activated, as if the anticipation was indeed actually miserable. Subjects also exhibit related behaviors, such as not collecting free information if it is likely to provide bad news.

These phenomena can be decision-theoretically rebadged by appealing to a psychologically rich utility model (Caplin and Leahy, 2001). The question for us is the psychological context of these utilities.

Three Pavlovian issues appear to be important. First, the activation of the primary pain system is consistent with a Pavlovian phenomenon called *stimulus substitution*, in which predictors of particular outcomes are treated in many respects like those outcomes themselves. Although the neural foundations of this are not clear, let alone its evolutionary rationale, it is an effect that is widely described, particularly in appetitive circumstances. For instance, the way that a pigeon treats a key which has a Pavlovian association with an appetitive outcome depends directly on whether it is food or water that is predicted. The pecks that result are recognizably associated with the specific outcome itself. The activation of the primary pain areas may arise through model-based stimulus substitution. If this then leads to an effective overcounting of the temporally distant shock, it can make the subject prefer the immediate one.

The other two Pavlovian effects are related to those discussed in the context of depressive realism. Not seeking information that is likely to be aversive is exactly akin to not exploring, or actually pruning, paths of thought that are likely to lead to negative outcomes. For dread itself, we can speculate as to the effects of the guaranteed prospect of a substantially delayed, future aversive outcome whose occurrence cannot be accurately predicted because of the inaccuracy in timing intervals (Gibbon *et al.*, 1997). This has both model-based and model-free consequences for the Pavlovian mechanism that creates optimism through pruning. From a model-based perspective, it creates a prior expectation of environments that are relatively unpleasant because they contain unpredictable aversive outcomes. Such environments are in general associated with larger average aversive values and so lead to Pavlovian avoidance (Huys and Dayan, 2008). From a model-free perspective, the persistent expectation of an aversive outcome might set a baseline level for the Pavlovian mechanism that prunes negative lines of thought. Since this baseline would be substantially more negative than usual, it would permit substantially more negative paths than normal to be explored, and therefore lead to net aversion.

AVERSIVELY MOTIVATED BEHAVIOR

We have so far used the analysis of the architecture of choice to highlight how Pavlovian predictions of aversive outcomes can lead to aberrant influences over instrumental choices in a wide variety of circumstances. However, there is an important instrumental component to aversive behavior too. Despite the apparent lack of current neuroeconomic interest in the topic, we will discuss avoidance, which is perhaps the most important such paradigm.

In an avoidance experiment, animals (or humans) learn actions that reliably lead to their avoiding incurring losses or pains. Typically, an animal receives a warning stimulus (such as a tone or light) that precedes delivery of an aversive stimulus, such as prolonged electrification of the floor of one compartment of the experimental apparatus. At first, the individual responds only during the aversive stimulus – for instance, escaping the shock by jumping into a neighboring compartment. Conventionally, the warning stimulus will be extinguished following this escape response. After several presentations, the escape response is executed more quickly, and eventually the individual learns to jump when observing the warning stimulus (again with the effect of turning off this stimulus), thus completely avoiding the shock.

Consideration of the problems that must be solved in avoidance hints that such behavior may not be straightforward. For instance, how are successful avoidance actions reinforced, if by definition they lead to no outcome? (How) does a subject ever realize that the threat is gone, if it is never sampled?

Mowrer famously suggested that learning to avoid involves two processes: predicting the threat, and learning to escape from the predictor (Mowrer, 1947). These processes, proposed respectively to be under Pavlovian and instrumental control, comprise two-factor theory, which in one form or another has survived well over the past decades. Although there are many unanswered questions about precisely how the different action systems are orchestrated in different avoidance situations, some key facts are well supported.

In particular, Pavlovian mechanisms play a critical (and multifarious) role in avoidance, and indeed Pavlovian responses to the warning stimulus alone are often capable of implementing successful avoidance. For example, jumping out of an electrified chamber, blinking in anticipation of an eye-puff, and leg flexion to an electric foot plate can all completely remove an aversive stimulus, without any need for an instrumental component. That they do pays tribute to their evolutionary provenance, and led some to question the involvement of instrumental responses at all (Mackintosh, 1983 for review). The latter is implied by the experimenter-controlled arbitrariness of the required avoidance actions – although more arbitrary ones are slower to learn (Biederman *et al.*, 1964; Riess, 1971; Ferrari *et al.*, 1973; Hineline, 1977).

Further, there is good evidence that the safety state that arises from successful avoidance acts as a Pavlovian aversive inhibitor (Rescorla, 1969; Weisman and Litner, 1969a; Morris, 1975; Dinsmoor, 2001; Candido *et al.*, 2004) – that is, a state that predicts the absence of otherwise expected punishment. Importantly, as mentioned above, the values of aversive inhibitors at least partly share a common representation with those of appetitive excitators, i.e. predictors of rewards, as is demonstrated by their ability to affect subsequent learning in appetitive domains (a phenomenon knows as *transreinforcer blocking*). That the safety state plays an important role in control is suggested by the fact that avoidance responses continue long after the Pavlovian aversive responses to the discriminative stimulus have extinguished, as they will of course do if avoidance is successful (Weisman and Litner, 1969b).

This places in the spotlight the role of the value attached to the warning stimulus (Kamin *et al.*, 1963; Biederman, 1968; De Villiers, 1974; Bersh and Lambert, 1975; Overmier *et al.*, 1971; Starr and Mineka, 1977; Mineka and Gino, 1980). On one hand, it has the power to initiate Pavlovian preparatory responses. It is also known to be able to suppress appetitive instrumental behavior, in a similar fashion to conditioned suppression by an aversive Pavlovian predictor. On the other hand, it has the instrumental power to initiate an appropriate avoidance response.

The dissociation of components in avoidance is supported by neural data. For instance, selective lesions of the central amygdala selectively impair conditioned suppression (aversive PIT) (Killcross *et al.*, 1997). Further, neuroleptics, which are dopamine antagonists, interfere with learning avoidance responses, but not acquisition of instrumental escape responses (Cook and Catania, 1964). This effect is of particular interest, since it suggests that it may only be the dopaminergically-reported appetitive outcome of reaching the safety state that can control instrumental learning of the avoidance response, as if the reduction of the aversive prediction itself is insufficient. This would be a very strange asymmetry between appetitive and aversive systems, and merits closer investigation.

In human studies, in support of the role of appetitive pathways, dorsal striatum and ventromedial prefrontal cortex display reward-signed activities during avoidance (Kim *et al.*, 2006; Pessiglione *et al.*, 2006). Furthermore, they do so in a manner predicted by reinforcement learning models.

There are known to be model-based components to avoidance learning. As discussed earlier in the chapter, one signature of this is the immediate sensitivity of actions to changes in the state of the subject that change the values of outcomes. An example of this outcome-sensitivity is an experiment that manipulated body temperature. Henderson and Graham (1979) trained rats to avoid a heat source when the rats were themselves hot. They then made the animals cold before testing them, and found that avoidance was attenuated, provided the rats had had the opportunity to experience the heat source in their new, cold state, thereby learning that it was rewarding. Selective lesions that dissociate goal-directed and habit-based components of the avoidance action are, however, currently lacking.

Sampling biases also pose a particular problem for aversive learning, since subjects will be unwilling to try options with aversive consequences in order to hone their behavior (Denrell and March, 2001). In fact, the sloth of extinction in avoidance is an example of this – if successful avoidance becomes reliably executed, how will the organism know if the threat has disappeared (termed the "hot stove effect" in economics)?. This contrasts with the appetitive case in which extinction is immediately frustrating. Pavlovian withdrawal will also severely hinder learning actions that lead to small, immediate, losses, but large, delayed, gains.

Of course, unnecessary avoidance is only economically problematic if there is some non-negligible cost to performing the action or if, unbeknownst to the organism, the action now leads to rewards. The problem of correctly navigating this issue is an example of the famous exploration–exploitation dilemma, which is raised in Chapter 24. Briefly, the battle is between exploiting existing knowledge, namely the lack of punishment that evidently ensues from performing the avoidance action, and exploring the possibility that the environment has changed such that the punishment is no longer present. The optimal solution to this dilemma is radically computationally intractable, since it depends on calculations associated with the uncertainties of unknown change. One conventional approximate approach is to behave non-deterministically, thus constantly sampling apparently lower-valued options stochastically. Another (sometimes more proficient) alternative is specifically to target actions whose consequences are more uncertain, as in uncertainty "bonus" schemes. The effect of these, in either appetitive or aversive domains, is to make subjects less risk- (and indeed ambiguity-) averse.

In sum, there is a substantial, subtle, literature on learned avoidance showing a range of intricate effects. Presently, little of this has had an impact in neuroeconomic paradigms, but it is a ripe area for exploration.

CONCLUSIONS

Aversion is not merely reward viewed through a looking glass. As we have reviewed here, aversion poses its own range of critical representational and learning phenomena, and exerts an important influence over a wealth of ecologic and economic tasks. We have focused on just a few of these – the substantial Pavlovian effects over experimental-, behavioral- and neuro-economic constructs, and the intricate complexities of avoidance learning – but there are also many other central issues that are being actively explored. From an economic perspective, much flows from the basic finding that mere monetary losses act in a very wide range of ways like real pains, thus allowing direct generalization from (and indeed to) an extensive psychological and neural literature.

Opponency has been a central concept in this chapter, as indeed it has over a wealth of psychological investigations. Unfortunately, although it is relatively uncontroversial that one of the opponents is dopamine, the identity, nature, and even exact functional role of the other is much less clear. We and others have argued in favor of the involvement of 5-HT, however, this is not yet totally accepted. Further, whether 5-HT, or the opponent, reports *all* punishments, or, for instance, only those punishments that are uncontrollable, or something else, is not yet evident.

Aversion is critical, pervasive, and interesting. Most relevantly, it is in clear need of the theoretical sophistication of neuroeconomic methods and analyses that are evidently on offer.

Acknowledgments

We are very grateful to our collaborators for discussions and ideas in these studies: Richard Bentall, Y.-Lan Boureau, Nathaniel Daw, Ray Dolan, Quentin Huys, Michael Moutoussis, Yael Niv, and John O'Doherty. We also thank Paul Glimcher and Antonio Rangel for comments on an earlier version of this chapter. Funding was from the Gatsby Charitable Foundation.

References

Abramson, L.Y., Metalsky, G.I., and Alloy, L.B. (1979). Judgment of contingency in depressed and nondepressed students: sadder but wiser? *J. Exp. Psychol. Gen.* 108, 441–485.

Adams, C.D. and Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Q. J. Exp. Psychol. B Comp. Physiol. Psychol.* 33, 109–121.

Balleine, B.W. and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419.

Balleine, B.W. and Killcross, S. (2006). Parallel incentive processing: an integrated view of amygdala function. *Trends Neurosci.* 29, 272–279.

Barto, A.G., Sutton, R.S., and Anderson, C.W. (1983). Neuronlike elements that can solve difficult learning problems. *IEEE Trans. Syst. Man Cybern.* 13, 834–846.

Barto, A.G., Sutton, R.S., and Watkins, C.J.C.H. (1990). Learning and sequential decision making. In: M. Gabriel and J. Moor (eds), *Learning and Computational Neuroscience: Foundations of Adaptive Networks*. Cambridge, MA: MIT Press, pp. 539–602.

Belova, M.A., Paton, J.J., Morrison, S.E., and Salzman, C.D. (2007). Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. *Neuron* 55, 970–984.

Bentham, J. (1817). *A Table of the Springs of Action, Showing the Several Species of Pleasures and Pains, of which Man's Nature is Susceptible*. London: R. & A. Taylor.

Bentham, J. (1823). *An Introduction to the Principles of Morals and Legislation*. London: T. Payne.

Berns, G.S., Chappelow, J., Cekic, M. *et al.* (2006). Neurobiological substrates of dread. *Science* 312, 754–758.

Bersh, P.J. and Lambert, J.V. (1975). Discriminative control of free-operant avoidance despite exposure to shock during stimulus correlated with nonreinforcement. *J. Exp. Anal. Behav.* 23, 111–120.

Bertsekas, D.P. (1995). *Dynamic Programming and Optimal Control*. Nashua, NH: Athena Scientific.

Biederman, G. (1968). Discriminated avoidance conditioning – Cs function during avoidance acquisition and maintenance. *Psychonomic Sci.* 10, 23–27.

Biederman, G., D'Amato, M.R., and Keller, D. (1964). Facilitation of discriminated avoidance learning by dissociation of CS and manipulandum. *Psychonomic Sci.* 1, 229–230.

Blanchard, R.J. and Blanchard, D.C. (1990). Anti-predator defense as models of animal fear and anxiety. In: P.F. Brain, R.J. Blanchard, and S. Parmigiani (eds), *Fear and Defense*. London: Harwood Academic, pp. 89–108.

Boyd, R. and Richerson, P.J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology Sociobiol.* 13, 171–195.

Breland, K. and Breland, M. (1961). The misbehavior of organisms. *Am. Psychologist* 16, 681–684.

Brown, P.L. and Jenkins, H.M. (1968). Auto-shaping of the pigeon's key-peck. *J. Exp. Anal. Behav.* 11, 1–8.

Camerer, C. (1995). Individual decision making. In: J.H. Kagel and A.E. Roth (eds), *The Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press.

Candido, A., Gonzalez, F., and de Brugada, I. (2004). Safety signals from avoidance learning but not from yoked classical conditioning training pass both summation and retardation tests for inhibition. *Behavioral Proc.* 66, 153–160.

Caplin, A. and Leahy, J. (2001). Psychological expected utility theory and anticipatory feelings. *Q. J. Economics* 116, 55–79.

Cardinal, R.N. (2006). Neural systems implicated in delayed and probabilistic reinforcement. *Neural Networks* 19, 1277–1301.

Cardinal, R.N., Parkinson, J.A., Hall, J., and Everitt, B.J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci. Biobehav. Rev.* 26, 321–352.

Cardinal, R.N., Winstanley, C.A., Robbins, T.W., and Everitt, B.J. (2004). Limbic corticostriatal systems and delayed reinforcement. *Adolesc. Brain Dev. Vuln. Opp.* 1021, 33–50.

Carter, R.M., O'Doherty, J.P., Seymour, B. *et al*. (2006). Contingency awareness in human aversive conditioning involves the middle frontal gyrus. *NeuroImage* 29, 1007–1012.

Chamberlain, S.R. and Sahakian, B.J. (2007). The neuropsychiatry of impulsivity. *Curr. Opin. Psychiatry* 20, 255–261.

Clutton-Brock, T.H. and Parker, G.A. (1995). Punishment in animal societies. *Nature* 373, 209–216.

Cook, L. and Catania, A.C. (1964). Effects of drugs on avoidance and escape behavior. *Fed. Proc.* 23, 818–835.

Corbit, L.H. and Balleine, B.W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *J. Neurosci.* 25, 962–970.

Corbit, L.H., Muir, J.L., and Balleine, B.W. (2001). The role of the nucleus accumbens in instrumental conditioning: evidence of a functional dissociation between accumbens core and shell. *J. Neurosci.* 21, 3251–3260.

Craig, A.D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3, 655–666.

Daw, N.D., Kakade, S., and Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Networks* 15, 603–616.

Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711.

Dayan, P. (2008). The role of value systems in decision making. In: C. Engel and W. Singer (eds), *Better Than Conscious? Implications for Performance and Institutional Analysis*. Cambridge, MA: MIT Press, Strungmann Forum Report.

Dayan, P. and Balleine, B.W. (2002). Reward, motivation, and reinforcement learning. *Neuron* 36, 285–298.

Dayan, P., Niv, Y., Seymour, B., and Daw, D. (2006). The misbehavior of value and the discipline of the will. *Neural Networks* 19, 1153–1160.

Delgado, M., Labouliere, C., and Phelps, E. (2006). Fear of losing money? Aversive conditioning with secondary reinforcers. *Social Cogn. Affect. Neurosci.* 1, 250–259.

De Martino, B., Kumaran, D., Seymour, B., and Dolan, R.J. (2006). Frames, biases, and rational decision-making in the human brain. *Science* 313, 684–687.

Denrell, J. and March, J.G. (2001). Adaptation as information restriction: the hot stove effect. *Organization Sci.* 12, 523–538.

De Villiers, P.A. (1974). The law of effect and avoidance: a quantitative relationship between response rate and shock-frequency reduction. *J. Exp. Anal. Behav.* 21, 223–235.

De Waal, F.B.M. (1998). *Chimpanzee Politics: Power and Sex Among Apes*. Baltimore, MD: Johns Hopkins University Press.

Dickinson, A. (1980). *Contemporary Animal Learning Theory*. Cambridge: Cambridge University Press.

Dickinson, A. and Balleine, B.W. (2002). The role of learning in motivation. In: C.R. Gallistel (ed.), *Learning, Motivation and Emotion, Vol. 3 of Steven's Handbook of Experimental Psychology*, 3rd edn. New York: John Wiley & Sons, pp. 497–533.

Dickinson, A. and Dearing, M.F. (1979). Appetitive-aversive interactions and inhibitory processes. In: A. Dickinson and R.A. Boakes (eds), *Mechanisms of Learning and Motivation*. Hillsdale, NJ: Erlbaum, pp. 203–231.

Dinsmoor, J.A. (2001). Stimuli inevitably generated by behavior that avoids electric shock are inherently reinforcing. *J. Exp. Anal. Behav.* 75, 311–333.

Estes, W.K. (1948). Discriminative conditioning. 2. Effects of A Pavlovian conditioned stimulus upon a subsequently established operant response. *J. Exp. Psychol.* 38, 173–177.

Estes, W.K. and Skinner, B.F. (1941). Some quantitative properties of anxiety. *J. Exp. Psychol.* 29, 390–400.

Everitt, B.J. and Robbins, T.W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* 8, 1481–1489.

Fanselow, M.S. (1994). Neural organization of the defensive behavior system responsible for fear. *Psychonomic Bull. Rev.* 1, 429–438.

Fehr, E. and Gachter, S. (2002). Altruistic punishment in humans. *Nature* 415, 137–140.

Fendt, M. and Fanselow, M.S. (1999). The neuroanatomical and neurochemical basis of conditioned fear. *Neurosci. Biobehav. Rev.* 23, 743–760.

Ferrari, E.A., Todorov, J.C., and Graeff, F.G. (1973). Nondiscriminated avoidance of shock by pigeons pecking a key. *J. Exp. Anal. Behav.* 19, 211–218.

Fields, H. (2004). State-dependent opioid control of pain. *Nat. Rev. Neurosci.* 5, 565–575.

Gibbon, J., Malapani, C., Dale, C.L., and Gallistel, C.R. (1997). Toward a neurobiology of temporal cognition: advances and challenges. *Curr. Opin. Neurobiol.* 7, 170–184.

Graeff, F.G. (2004). Serotonin, the periaqueductal gray and panic. *Neurosci. Biobehav. Rev.* 28, 239–259.

Graeff, F.G., Guimaraes, F.S., DeAndrade, T.G.C.S., and Deakin, J.F.W. (1996). Role of 5-HT in stress, anxiety, and depression. *Pharmacol. Biochem. Behav.* 54, 129–141.

Grossberg, S. (1984). Some normal and abnormal behavioral syndromes due to transmitter gating of opponent processes. *Biol. Psychiatry* 19, 1075–1118.

Henderson, R.W. and Graham, J. (1979). Avoidance of heat by rats – effects of thermal context on rapidity of extinction. *Learning Motiv.* 10, 351–363.

Hineline, P.N. (1977). Negative reinforcement and avoidance. In: W.K. Honig and J.E.R. Staddon (eds), *Handbook of Operant Behavior*. Englewood Cliffs, NJ: Prentice Hall, pp. 364–414.

Horvitz, J.C. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* 96, 651–656.

Huys, Q. and Dayan, P. (2008). A Bayesian formulation of behavioral control. *Cognition*, (in press).

Ikemoto, S. and Panksepp, J. (1999). The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Res. Brain Res. Rev.* 31, 6–41.

Jensen, J., McIntosh, A.R., Crawley, A.P. *et al*. (2003). Direct activation of the ventral striatum in anticipation of aversive stimuli. *Neuron* 40, 1251–1257.

Jensen, J., Smith, A.J., Willeit, M. *et al*. (2007). Separate brain regions code for salience vs. valence during reward prediction in humans. *Hum. Brain Mapp.* 28, 294–302.

Julius, D. and Basbaum, A.I. (2001). Molecular mechanisms of nociception. *Nature* 413, 203–210.

Kahneman, D. and Tversky, A. (2000). *Choices, Values, and Frames*. Cambridge: Cambridge University Press.

Kamin, L.J., Black, A.H., and Brimer, C.J. (1963). Conditioned suppression as a monitor of fear of Cs in course of avoidance training. *J. Comp. Physiol. Psychol.* 56, 497.

Killcross, S., Robbins, T.W., and Everitt, B.J. (1997). Different types of fear-conditioned behavior mediated by separate nuclei within amygdala. *Nature* 388, 377–380.

Kim, H., Shimojo, S., and O'Doherty, J.P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PloS Biol.* 4, 1453–1461.

Knutson, B., Rick, S., Wirnmer, G.E. *et al*. (2007). Neural predictors of purchases. *Neuron* 53, 147–156.

Koechlin, E., Ody, C., and Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science* 302, 1181–1185.

Konorski, J. (1967). *Integrative Activity of the Brain: An Interdisciplinary Approach*. Chicago, IL: University of Chicago Press.

LeDoux, J.E. (2000). Emotion circuits in the brain. *Annu. Rev. Neurosci.* 23, 155–184.

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *Economic J.* 97, 666–684.

Loewenstein, G. (2006). The pleasures and pains of information. *Science* 312, 704–706.

Lovibond, P.F. (1983). Facilitation of instrumental behavior by a Pavlovian appetitive conditioned stimulus. *J. Exp. Psychol. Animal Behav. Proc.* 9, 225–247.

Lumb, B.M. (2002). Inescapable and escapable pain is represented in distinct hypothalamic-midbrain circuits: specific roles for A delta- and C-nociceptors. *Exp. Physiol.* 87, 281–286.

Mackintosh, N.J. (1983). *Conditioning and associative learning*. New York, NY: Oxford University Press.

Maier, S.F. and Seligman, M.E.P. (1976). Learned helplessness – theory and evidence. *J. Exp. Psychol. Gen.* 105, 3–46.

Mangel, M. and Clark, C.W. (1988). *Dynamic Modelling in Behavioral Ecology*. Princeton, NJ: Princeton University Press.

Maren, S. and Quirk, G.J. (2004). Neuronal signalling of fear memory. *Nat. Rev. Neurosci.* 5, 844–852.

Melvin, K.B. and Anson, J.E. (1969). Facilitative effects of punishment on aggressive behavior in Siamese fighting fish. *Psychonomic Sci.* 14, 89–90.

Mineka, S. and Gino, A. (1980). Dissociation between conditioned emotional response and extended avoidance performance. *Learning Motiv.* 11, 476–502.

Mobbs, D., Petrovic, P., Marchant, J.L. *et al*. (2007). When fear is near: threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science* 317, 1079–1083.

Montague, P.R., Dayan, P., and Sejnowski, T.J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.

Morris, R.G.M. (1975). Preconditioning of reinforcing properties to an exteroceptive feedback stimulus. *Learning Motiv.* 6, 289–298.

Morse, W.H., Mead, R.N., and Kelleher, R.T. (1967). Modulation of elicited behavior by a fixed-interval schedule of electric shock presentation. *Science* 157, 215–217.

Mowrer, O.H. (1947). On the dual nature of learning: a re-interpretation of "conditioning" and problem-solving. *Harv. Educat. Rev.* 17, 102–148.

Nakahara, H., Itoh, H., Kawagoe, R. *et al*. (2004). Dopamine neurons can represent context-dependent prediction error. *Neuron* 41, 269–280.

Niv, Y., Joel, D., and Dayan, P. (2006). A normative perspective on motivation. *Trends Cogn. Sci.* 10, 375–381.

Niv, Y., Daw, N.D., Joel, D., and Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology* 191, 507–520.

Overmier, J.B., Bull, J.A., and Trapold, M.A. (1971). Discriminative cue properties of different fears and their role in response selection in dogs. *J. Comp. Physiol. Psychol.* 76, 478–482.

Paton, J.J., Belova, M.A., Morrison, S.E., and Salzman, C.D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439, 865–870.

Pessiglione, M., Seymour, B., Flandin, G. *et al*. (2006). Dopamine-dependent prediction errors underpin reward-seeking behavior in humans. *Nature* 442, 1042–1045.

Price, D.D. (1999). *Psychological Mechanisms of Pain and Analgesia*. Seattle, WA: IASP Press.

Puterman, M. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY: John Wiley & Sons, Inc.

Rescorla, R.A. (1969). Establishment of a positive reinforcer through contrast with shock. *J. Comp. Physiol. Psychol.* 67, 260–263.

Riess, D. (1971). Shuttleboxes, Skinner boxes, and Sidman aoidance in rats: acquisition and terminal performance as a function of response topography. *Psychonomic Sci.* 25, 283–286.

Sachs, B.D. and Barfield, R.J. (1974). Copulatory behavior of male rats given intermittent electric shocks: theoretical implications. *J. Comp. Physiol. Psychol.* 86, 607–615.

Satoh, T., Nakai, S., Sato, T., and Kimura, M. (2003). Correlated coding of motivation and outcome of decision by dopamine neurons. *J. Neurosci.* 23, 9913–9923.

Schoenbaum, G. and Setlow, B. (2003). Lesions of nucleus accumbens disrupt learning about aversive outcomes. *J. Neurosci.* 23, 9833–9841.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Setlow, B., Schoenbaum, G., and Gallagher, M. (2003). Neural encoding in ventral striatum during olfactory discrimination learning. *Neuron* 38, 625–636.

Seymour, B., O'Doherty, J.P., Dayan, P. *et al*. (2004). Temporal difference models describe higher-order learning in humans. *Nature* 429, 664–667.

Seymour, B., O'Doherty, J.P., Koltzenburg, M. *et al*. (2005). Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat. Neurosci.* 8, 1234–1240.

Seymour, B., Daw, N., Dayan, P. *et al*. (2007a). Differential encoding of losses and gains in the human striatum. *J. Neurosci.* 27, 4826–4831.

Seymour, B., Singer, T., and Dolan, R. (2007b). The neurobiology of punishment. *Nat. Rev. Neurosci.* 8, 300–311.

Shidara, M., Mizuhiki, T., and Richmond, B.J. (2005). Neuronal firing in anterior cingulate neurons changes modes across trials in single states of multitrial reward schedules. *Exp. Brain Res.* 163, 242–245.

Solomon, R.L. and Corbit, J.D. (1974). An opponent-process theory of motivation. I. Temporal dynamics of affect. *Psychol. Rev.* 81, 119–145.

Starr, M.D. and Mineka, S. (1977). Determinants of fear over course of avoidance-learning. *Learning Motiv.* 8, 332–350.

Stevens, J.R. (2004). The selfish nature of generosity: harassment and food sharing in primates. *Proc. Biol. Sci.* 271, 451–456.

Sugase-Miyamoto, Y. and Richmond, B.J. (2005). Neuronal signals in the monkey basolateral amygdala during reward schedules. *J. Neurosci.* 25, 11071–11083.

Sutton, R.S. and Barto, A.G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychol. Rev.* 88, 135–170.

Sutton, R.S. and Barto, A.G. (1998). *Reinforcement Learning. An Introduction*. Cambridge, MA: MIT Press.

Thomson, J.J. (1986). *Rights, Restitution and Risk*. Cambridge, MA: Harvard University Press, pp. 94–116.

Tolman, E.C. (1932). *Purposive Behavior in Animals and Men*. New York, NY: Century.

Tversky, A. and Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science* 211, 453–458.

Ulrich, R.E. and Azrin, N.H. (1962). Reflexive fighting in response to aversive stimulation. *J. Exp. Anal. Behav.* 5, 511–520.

Ursu, S. and Carter, C.S. (2005). Outcome representations, counterfactual comparisons and the human orbitofrontal cortex: implications for neuroimaging studies of decision-making. *Brain Res. Cogn. Brain Res.* 23, 51–60.

Watkins, C.J.C.H. and Dayan, P. (1992). Q-Learning. *Machine Learning* 8, 279–292.

Weisman, R.G. and Litner, J.S. (1969a). Positive conditioned reinforcement of Sidman avoidance behavior in rats. *J. Comp. Physiol. Psychol.* 68, 597–603.

Weisman, R.G. and Litner, J.S. (1969b). The course of Pavlovian excitation and inhibition of fear in rats. *J. Comp. Physiol. Psychol.* 69, 667–672.

Williams, D.R. and Williams, H. (1969). Auto-maintenance in pigeon – sustained pecking despite contingent non-reinforcement. *J. Exp. Anal. Behav.* 12, 511–520.

Wilson, D.I. and Bowman, E.M. (2005). Rat nucleus accumbens neurons predominantly respond to the outcome-related properties of conditioned stimuli rather than their behavioral-switching properties. *J. Neurophysiol.* 94, 49–61.

Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *J. Pers. Social Psychol.* 51, 110–116.

Yin, H.H. and Knowlton, B.J. (2006). The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* 7, 464–476.

Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action–outcome contingency in instrumental conditioning. *Behav. Brain Res.* 166, 189–196.

Yoshida, W. and Ishii, S. (2006). Resolution of uncertainty in prefrontal cortex. *Neuron* 50, 781–789.

CHAPTER

13

Behavioral Game Theory and the Neural Basis of Strategic Choice

Colin F. Camerer

OUTLINE

GAME THEORY

Game theory is a very general language for modeling choices by purposive agents in which the actions of other agents can affect each player's outcomes. The elements of a game are players, strategies, information, a structure or "game form" (e.g., who chooses when), outcomes which result from all players' strategy choices and information, and preferences over those outcomes. The generality of this language makes it applicable to many levels of analysis (from biology to international politics). Examples include evolutionary adaptation of genes competing on the basis of fitness, face-to-face bargaining, competition and collusion among firms, behavior in rule-governed games like chess and poker which are computationally demanding, inferring something about a person's preferences or future intentions from their actions, delicate diplomatic bargaining in which language has multiple meaning to different audiences, and designing rules to raise the most revenue from government auctions of scarce resources.

The aim of this chapter is to introduce some of the essential components of game theory to neuroscientists (for more details, see Chapter 5 in this volume), and to summarize some emerging regularities from psychological and neural analysis which suggest ways to put the theory of behavior in games on a biological basis.

Game theory is useful in two ways. Simply specifying the details of a strategic interaction and matching it to familiar categories of games (such as the prisoners' dilemma game) is helpful as a language to describe a situation and point out its crucial features (Aumann, 1985). Classification of this sort can be helpful even if there is no deep mathematical analysis of what players are likely to do.

Most of game theory, however, analyzes what players are likely to do once a game is fully specified mathematically. Analytical game theory assumes players choose strategies which maximize utility of game outcomes given their beliefs about what others players will do. This means that the most challenging question is often how beliefs are formed. Most theories assume that beliefs are derived from some kind of analysis of what other players are likely to do, given the economic structure of the game. In equilibrium analysis, these beliefs about others are assumed to be correct, which solves the problem of how to specify reasonable beliefs by equating them with choices. While analytical game theory has proved enormously powerful, there are two shortcomings of its tools that limit its use as a complete model of behavior by people (and other levels of players).

First, many of the games that occur naturally in social life are so complex that it is unlikely that players instantaneously form accurate beliefs about what others would do and therefore can choose equilibrium strategies. It is therefore useful to consider what strategies might be chosen by players with bounded rationality, or when there is learning from repeated play[1].

Second, in empirical work, only received (or anticipated) payoffs are easily measured (e.g. prices and valuations in auctions, the outcome of a union–management wage bargain, or currency paid in an experiment). But game theory takes as its primitives the *preferences* players have for the received payoffs of all players (utilities), and preferences are generally taken to be most clearly revealed by actual choices (see Chapters 3 and 4 in this volume). Inferring from strategic choices alone both the beliefs players have about choices of others, and their preferences for outcomes which result from mutual choices, is therefore especially challenging. One shortcut is to have a theory of theory of social preferences – how measured payoffs for all players from an outcome determine players' utility evaluations of that outcome – in order to make predictions. Emerging concepts of social preference and their neural correlates are reviewed by Fehr (Chapter 15) and Camerer (Chapter 13).

Hundreds of experiments show that analytical game theory sometimes explains behavior surprisingly well, and is sometimes badly rejected by behavioral and process data (Camerer, 2003). This wide range of data – when game theory works well and badly – can be used to create a more general theory which approximately matches the standard theory when it is accurate, and can explain the cases in which it is badly rejected. This chapter describes an emerging approach called "behavioral game theory," which generalizes analytical game theory to explain experimentally-observed violations by incorporating bounds on rationality in a formal way.

Like analytical game theory, behavioral game theory is efficiently honed by laboratory regularity because the structure of the game and resulting payoffs can be carefully controlled in the lab (in field applications it is usually hard to know what game the players think they are playing). However, behavioral game theory is ultimately aimed at practical questions like how workers react to employment terms, the evolution of Internet market institutions for centralized trading (including reputational systems), the design of auctions and contracts, explaining animal behavior, and players "teaching" other players who learn what to expect (such as firms intimidating competitors or building trust in strategic alliances, or diplomats threatening and cajoling).

BEHAVIORAL GAME THEORY

Behavioral game theory is explicitly meant to predict how humans (and perhaps firms and other collective entities) behave. It has four components: representation, social preferences over outcomes, initial conditions, and learning.

Representation

How is a game perceived or mentally represented? Often the game players perceive may be an incomplete representation of the true game, or some elements of the game may be ignored to reduce computationally complexity. This topic has been studied very little, however (Camerer, 1998).

One example is multi-stage alternating-offer bargaining. In this game, agents bargain over a sum of money, and alternate offers about how to divide the sum. If an offer is rejected, the available money shrinks (representing the loss of value from delay). The game ends when an offer is accepted. One version of the game that has been studied experimentally has three stages, with sums varied randomly around $5,

---

[1] A different approach, "evolutionary game theory", assumes that agents in a population play fixed strategies, but population pressure adjusts the statistical mixture of strategies across the population – i.e. the "market share" of each strategy in the population – so that successful strategies are reproduced more frequently.

$2.50, and $1.25 (if the last offer is rejected, both players get nothing). If players are self-interested and plan ahead, the prediction of game theory is that the player who makes the first offer should offer $1.26 and the other player will accept it[2]. Empirically, players offer more than predicted, around $2.10, and much lower offers are often rejected.

One possible explanation is that players care about fairness. Beliefs are in equilibrium but they reject low offers because they prefer to get a larger share of a smaller sum of money. Another explanation is that players do not plan ahead. If they act as though the game will last only two periods, for example, then the equilibrium offer is $2.50; so the empirical average offer of $2.10 might reflect some mixture of playing the three-period game and playing a truncated two-period game.

Camerer *et al*. (1993) and Johnson *et al*. (2002) compared these two explanations by using a "Mouselab" system which masks the three varying dollar amounts that are available in the three rounds in opaque boxes (like in the game show *Jeopardy*!). Information is revealed when a computer mouse is moved into the box (and the box closes when the mouse is moved outside of it). They found that in about 10–20% of the trials subjects did not even bother to open the box showing the sum of money that would be available in the second or third stage. Their information look-up patterns are also correlated with offers they make (subjects who looked ahead further made lower offers). By directly measuring whether players are opening the value boxes, and how long those boxes are open, they could conclude that subjects were computing based on an attentionally limited representation of the game. Keep in mind that these games are simple and players are capable of perceiving the entire game[3]. Further work on limited representations could study more complicated games where truncation of representations is likely to be even more dramatic and insightful.

SOCIAL PREFERENCES OVER OUTCOMES

As noted above, when the payoffs in a game are measured, a theory of preferences over payoff distributions is needed to fully specify the game. This is a rich area of research discussed by Camerer (2003: Chapter 2), Fehr and Camerer (2007); see also Chapter 15 of this volume).

Initial Conditions or One-shot Play

Many games are only played once (a "one-shot" game), and in other cases an identical game is played repeatedly (a "repeated game"). In many games, it is not plausible that beliefs will be correct immediately in one-shot games or in the first period of a repeated game (as assumed by equilibrium models), without pre-play communication or some other special condition. Two types of theories of initial conditions have emerged: cognitive hierarchy (CH) theories of limits on strategic thinking; and theories which retain the equilibrium assumption of equilibrium beliefs but assume players make stochastic mistakes.

Cognitive hierarchy theories start with the presumption that iterated reasoning is limited in the human mind, and heterogeneous across players. Limits arise from evolutionary constraint in promoting high-level thinking, limits on working memory[4], and adaptive motives for overconfidence in judging one's relative skill (i.e., people may stop after some iteration because they think others must not have thought any further than they did).

Denote the probability of a step-k player i choosing strategy j by $P_k(s_i^j)$. The payoffs for player i if the other player (denoted $-i$) chooses s_{-i}^h are given by a payoff function of both strategies denoted by $p_i(s_i^j, s_{-i}^h)$. Assume a distribution $f(k)$ of k-step types. Zero-step players choose randomly (i.e., $P_0(s_i^j) = 1/n$ if there are n strategies). k-step players form a conditional belief $g_k(t)$ about the percentage of opponents who do

---

[2] Assuming players only care about their own payoffs, the prediction of game theory comes from forecasting what would happen at every future "subgame" and working backward ("backward induction"). In the third stage, player 1 should expect that an offer of $.01 will be accepted, leaving $1.24 for himself. Player 2 should anticipate this, and offer $1.25 to player 1 out of the total of $2.50 in the second stage (just a penny more than player 1 expects to get in the third stage), leaving $1.25 for himself (player 2) in the second stage. In the first round, player 1 should anticipate that player 2 expects to earn $1.25 in the second stage and offer $1.26.

[3] In one condition, subjects play against a computer algorithm which they know is optimized to earn the highest payoff, and expects that human players will do the same. At first, subjects' looking strategies and offers are similar to those when they play human opponents. However, when they are gently told that it might be useful to look at all three amounts available for bargaining and work backward ("backward induction"), subjects learn rapidly to play the optimal strategy and open all boxes.

[4] Devetag and Warglien (2003) show a correlation across subjects between working memory, as measured by digit span, and choices linked to the number of steps of thinking.

t steps of thinking. One specification is that k-step players guess the relative proportions of how much thinking other players do correctly, but they do not realize others might be doing k or more steps of thinking, that is

$$g_k(t) = f(t)/\left(\sum_{m=0}^{k-1} f(m)\right) \textit{ for } t = 0 \textit{ to } k - 1$$[5]

Given their beliefs, k-step players figure out what all lower-step types will do and form an expected payoff for strategy s_i^j

$$E_k(s_i^j) = \Sigma_t g_k(t) \Sigma_h P_t(s_j^h) p_i(s_i^j, s_j^h)$$

A level-k player responds with a logit (softmax) choice function

$$p_k(s_j^j) = \exp(\lambda E_k(s_i^j))/\Sigma_n \exp(\lambda E_k(s_i^n))$$

This model is easy to compute numerically because it uses a simple loop. Behavior of level-k players depends only on the behavior of lower-level players, which is computed earlier in the looping procedure. In contrast, equilibrium computation is often more difficult because it requires solving for a fixed-point vector of strategies which is a best response to itself.

A useful illustration of how the cognitive hierarchy approach can explain deviations from equilibrium analysis is the "p-beauty contest" game (Nagel, 1995; Ho *et al.*, 1998). In this game, several players choose a number in the interval [0, 100]. The average of the numbers is computed, and multiplied by a value p (in many studies, $p = 2/3$). The player whose number is closest to p times the average wins a fixed prize.

In equilibrium, by definition, players are never surprised what other players do. In the p-beauty contest game, this equilibrium condition implies that all players must be picking p times what others are choosing. This equilibrium condition only holds if everyone chooses 0 (the Nash Equilibrium, consistent with iterated dominance).

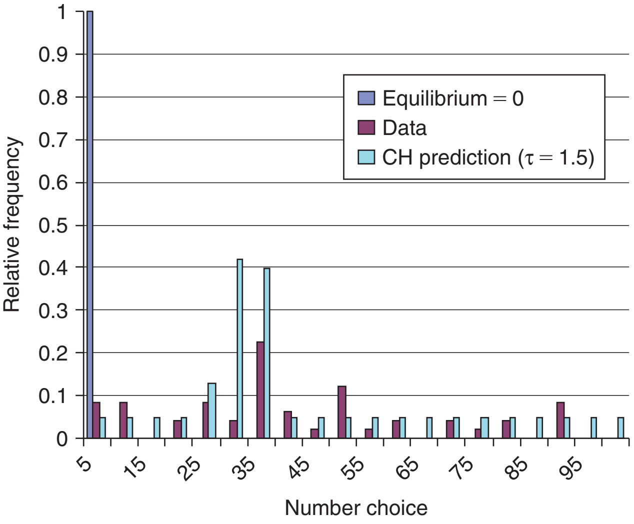


FIGURE 13.1 Data and predictions for .7 times the average game. Players choose numbers 0–100 simultaneously; the player closest to .7 times the average wins a fixed monetary prize. Data are closer to the prediction of a cognitive hierarchy (CH) model than to the equilibrium prediction of 0. Reproduced from Camerer and Fehr (2006), with permission.

Figure 13.1 shows data from a game with $p = .7$ and compares the Nash prediction (choosing 0) and the fit of a cognitive hierarchy model (Camerer *et al.*, 2004). In this game, some players choose numbers scattered from 0 to 100, many others choose p times 50 (the average if others are expected to choose randomly), and others choose p^2 times 50. Groups of subjects with high analytical skill and training in game theory do choose lower numbers (although 0 is still rarely chosen). When the game is played repeatedly with a fixed group of players who learn the winning number, number choices do converge toward zero – a reminder that equilibrium concepts can reliably predict where an adaptive learning process leads. Costa-Gomes *et al.* (2001), Camerer *et al.* (2004), Costa-Gomes and Crawford (2006) and earlier studies show how these cognitive hierarchy theories can fit experimental data from a wide variety of games, with similar thinking-step parameters across games. Hedden and Zhang (2002) show similar evidence of limited thinking from cognitive experiments.

The cognitive hierarchy theories deliberately allow for the possibility that some players do not correctly guess what others will do. Another approach, called "quantal response" equilibrium (QRE), retains the assumption that each player's beliefs are statistically correct, but uses a softmax choice function so that choices are not always payoff-maximizing best-responses. (That is, players can make mistakes in strategy choice, but large mistakes are rarer than small mistakes.) The expected payoff of player i's strategy s_i^h is $E(s_i^h) = \Sigma_k P_{-i}(s_j^k) p_i(s_i^h, s_{-i}^k)$, and choice probabilities are given by a softmax function (see above). QRE fits a wide variety of data better than do Nash predictions

---

[5] The simpler specification $g_k(h) = 1$ for $h = k - 1$ (k-steppers believe all others do exactly one less step) is often more tractable and is also widely used; see Camerer *et al.*, 2004; see also Nagel, 1995; Stahl and Wilson, 1995; Costa-Gomes *et al.*, 2001; Costa-Gomes and Crawford, 2006; Wang *et al.*, 2006.

TABLE 13.1 “Work–shirk” game payoffs (Dorris and Glimcher, 2004)

		Employer	
		Inspect	**Don’t inspect**
Worker	Work	.5, 2–I	.5, 2
	Shirk	0, 1–I	1, 0

Note: Employer mixed-strategy equilibrium probability is (.5, inspect; .5, don’t inspect). Worker mixed-strategy equilibrium probability is (I, shirk; 1–I, work).

TABLE 13.2 Variation in equilibrium shirking rates, cognitive hierarchy prediction, and actual human and monkey shirk rates in the work–shirk game

	Inspection cost I			
	.1	**.3**	**.7**	**.9**
Equilibrium p(shirk)	.10	.30	.70	.90
CH prediction ($\tau = 1.5$)	.28	.28	.72	.72
Human data	.29	.48	.69	.83
Monkey data	.30	.42	.64	.78

Data from Dorris and Glimcher (2004) (Table 13.1). CH predictions from online calculator at http://groups.haas.berkeley.edu/simulations/ch/default.asp.

(McKelvey and Palfrey, 1995, 1998; Goeree and Holt, 2001)[6].

The essential elements of CH and QRE can also be synthesized into a more general approach (Camerer *et al.*, 2008), although each of the simpler components fits a wide range of games about as well as a does more general hybrid model.

One goal of CH and QRE is to explain within a single model why behavior is far from equilibrium in some games (like the p-beauty contest) and remarkably close to equilibrium in others. An example is games with mixed equilibrium. In a mixed equilibrium, a player’s equilibrium strategy mixes probability across different strategies (that is, there is no combination of strategies played for sure – “pure strategies”– which is an equilibrium).

One game that has been studied relatively frequently in neuroeconomics, which only has a mixed equilibrium, is the “work or shirk” inspection game shown in Table 13.1 (Dorris and Glimcher, 2004). The economic story surrounding the game is that a lazy worker prefers not to work, but an employer knows this and sometimes “inspects” the worker. There is no pure equilibrium, because the worker only works because of the fear of inspection, and the employer does not inspect all the time if the worker is expected to work. Instead, both players mix their strategies. For the Table 13.1 game payoffs, employers inspect half the time and workers shirk I% of the time (where I is the cost of inspection). This game is in a class called “asymmetric matching pennies,” because the worker prefers to match strategies on the diagonal (working if they are inspected and shirking if they aren’t) and the employer prefers to mismatch.

Empirically, in games with mixed equilibria the relative frequencies of strategies chosen in the first period are actually remarkably close to the predicted frequencies (see Camerer, 2003: Chapter 3) although they are regressive: That is, actual play of strategies predicted to be rare (common) is too high (too low). Table 13.2 illustrates results from human and monkeys in Dorris and Glimcher (2004). The monkey and human data are very close. The CH prediction fits the data much better than the Nash prediction for I = .1, and is equally close for other values of I.

The CH model explains why choices are so close to the mixed equilibrium probabilities through the heterogeneity of players. Low-level players randomize, but higher-level players respond to expected randomization. The mixture across those level types tends to be close to the first-period data, but is closer to equal mixing than the equilibrium predictions (which typically fits data better).

CH models have also been applied in two field settings. In a Swedish lottery game called LUPI (Östling *et al.*, 2007), players chose integers from 1 to 99,999 and the *l*owest *u*nique *p*ositive *i*nteger wins (hence the name LUPI). About 50,000 people played the lottery each day. The (symmetric) Nash Equilibrium prediction is approximately equal choice of numbers from 1 to 5000, a sharp drop-off in choice from 5000 to 5500, and very few choices above 5500. This prediction is derived from extremely complicated algebra, using only the rules of the game, the range of numbers, and the number of players as inputs (with zero free parameters). Figure 13.2 shows that actual behavior in the first 7 days of play is surprisingly close to this prediction (shown by the dotted line). However, there is a clear tendency to choose too many low numbers,

---

[6] QRE also circumvents some technical limits of Nash Equilibrium. In Nash Equilibrium, players can put zero weight on the chance of a mistake or “tremble” by other players, which can lead to equilibria which are implausible because they rely on threats that would not be carried out at future steps of the game. In QRE, players always tremble and the degree of trembling in strategies is linked to expected payoff differences (*cf.* Myerson, 1986).

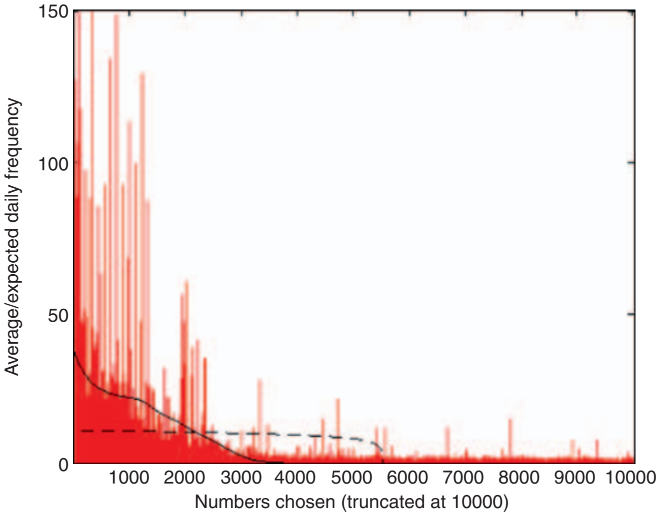


FIGURE 13.2 LUPI game results from the first week of Swedish lotteries. $n = 53{,}000$ players choose integers 1–99,999. The lowest unique integer wins a large prize (100,000 Swedish Krona, ~10,000 €). The symmetric equilibrium prediction is shown by the dotted line and CH best-fitting model is shown by the solid line ($\tau = 2.89$).

too few numbers from 2500–5000, and too many numbers above the drop-off at 5500. The cognitive hierarchy model (the solid line) explains these deviations reasonably well with a best-fitting value of $\tau = 2.98$, comparable to values from 1–2 which fit experimental data well. Brown *et al.* (2007) also use the CH model to explain why moviegoers seem to ignore the fact that movies which are not reviewed before they are released tend to be low in quality. This strategic naïveté leads to a box-office premium from withholding poor movies for review. Their analysis estimates a best-fitting $\tau = 1.26$ for moviegoer behavior, close to the LUPI game estimate and earlier lab estimates.

Learning

When a game is played repeatedly, agents can learn from the payoffs they get and from the strategies other players choose, and can also learn about what other players are likely to do. Many models of these learning processes have been proposed and tested on a wide variety of experimental games (see Camerer, 2003: Chapter 6). The general structure is that strategies have numerical attractions that are updated based on observation of payoffs and actions of other players. Attractions determine choice probabilities using a logit or comparable rule. The difference across models is how attractions are updated.

There are several important differences across models. Denote a strategy j's attraction for player i after period t by $A_i^j(t)$. In reinforcement learning, the attraction of the chosen strategy is updated by the received payoff

$$A_i^j(t) = \phi A_i^j(t-1) + (1-\phi)p_i(s_i^j, s_{-i}(t))$$

where $s_{-i}(t)$ is the strategy actually chosen by opponent $-i$ in period t and ϕ is a geometric decay. Note that this can be written as

$$A_i^j(t) = A_i^j(t-1) + (1-\phi)[p_i(s_i^j, s_{-i}(t)) - A_i^j(t-1)]$$

The payoff surprise $p_i(s_i^j, s_{-i}(t)) - A_i^j(t-1)$ is a prediction error – the difference between the received payoff and the previous attraction – so the learning rule is a form of temporal-difference learning rule (see Chapter 22 of this volume). A different approach is to update beliefs about what other players will choose, then use those new beliefs to update attractions – as in "fictitious play" learning, which keeps track of the fraction of previous choices by other players of each strategy (possibly geometrically-weighted to incorporate forgetting or perception of non-stationarity in opponent play).

Camerer and Ho (1999) noted that the reinforcement rule written above and fictitious play are both special cases of a more general "experience-weighted attraction" (EWA) family in which

$$A_i^j(t) = \{\phi N(t-1)A_i^j(t-1) + d(s_i^j, s_i(t)) * p_i(s_i^j, s_{-i}(t))\}/N(t)$$

where

$$N(t) = \phi(1-\kappa)N(t-1) + 1$$

is a cumulated weight of experience.

When $\kappa = 0$, the rule is a TD-like averaging rule. The weight on new payoff information is $1/(\phi N(t-1) + 1)$ which falls over time t as $N(t-1)$ grows, so that learning slows down. This algebraic form expresses a time-adjusted learning rate[7]. The key term is $d(s_i^j, s_i(t)) = \delta + (1-\delta)I(s_i^j, s_i(t))$, where $I(x, y)$ is an identity function which equals 1 if $x = y$ and 0 otherwise. This "imagination" weight is 1 for the chosen strategy and δ for unchosen strategies. When $\delta = 0$, the model reduces to reinforcement of the strategy that is actually played. When $\delta = 1$, it is mathematically equivalent to fictitious play; both payoffs from strategies that are actually played and "fictive"

---

[7] However, unlike the standard temporal difference rule, when $\kappa = 0$ the rule cumulates payoffs rather than averages them. This allows attractions to grow outside the bounds of payoffs which, in the softmax rule, means that probabilities can lock in sharply at extreme values of 0 or 1.

payoffs from unplayed strategies influence learning equally strongly. The insight here is that learning by updating beliefs about other players' choices (using fictitious play) is exactly the same, mathematically, as generalized reinforcement in which unchosen strategies are updated by the payoffs they would have created. In computer science terms, EWA represents a hybrid of model-free learning from choices and model-based learning (which uses information about unchosen strategy payoffs through a "model" which is the structure of the game).

Ho *et al.* (2007) propose and estimate a "self-tuning" version of EWA in which ϕ and δ are functions of experience (with $N(0) = 1$ and $\kappa = 0$ for simplicity), so that there are no free parameters except for the response sensitivity?. The function f is interpreted as a "change-detector" which adjusts the learning rate to environmental uncertainty. When another player's behavior is highly variable, or changes suddenly, ϕ falls so that more relative weight is placed on new payoff information. Behrens *et al.* (2007) have found neural evidence for such a learning-adjustment process in decision problems with non-stationary payoffs. Soltani *et al.* (2006) have simulated behavior of a similar "meta-learning" model which explores learning model parameters (Schweighofer and Doya, 2003) and shown that it fits some aspects of monkey behavior.

All these models are adaptive because they use only previous payoffs in the updating equation. However, eyetracking experiments show that players do look at payoffs of other players (Wang *et al.*, 2007) and are responsive to them. In the empirical game learning literature, learning rules that anticipate how other players might be learning are termed *sophisticated*. Stahl (2003) and Chong *et al.* (2006) proposed a sophisticated rule in which players believe that others are learning according to EWA and respond to expected payoffs based on that belief.

If players are playing together repeatedly, sophisticated players could also take account of their current actions when considering what other players will do in the future, a process called *strategic teaching*. Chong *et al.* (2006) showed evidence of strategic teaching in games based on on trust and entry deterrence.

PSYCHOLOGICAL AND NEURAL EVIDENCE

Behavioral game theory analyses of experimental data have proceeded along a parallel track with other psychological and neural studies, but the tracks have rarely met. This section mentions some types of neural activity which might be linked to behavioral game theory constructs in future research.

Theory of Mind (TOM) and Human-computer Differences

Theory-of-mind (TOM) refers to the capacity to make accurate judgments about the beliefs, desires, and intentions of other people, which are crucial inputs for appropriate social judgment and for social success (see Chapters 18 and 19 of this volume for further discussion of these issues across species). TOM is thought to be impaired in autism. It is widely thought that neural components of TOM include anterior and posterior cingulate, medial frontal cortex (Frith and Frith, 2006), paracingulate cortex, superior temporal sulcus (STS), and the temporal-parietal junction (TPJ). There is lively empirical debate about which of these regions are involved in different kinds of social reasoning and attribution. For example, Saxe and Powell (2006) argue that bilateral TPJ is unique for understanding another person's thoughts, and develops later in life, while mPFC is more useful for more general social understanding (e.g., sensations that other people feel).

If TOM is indeed a separate faculty, it certainly is necessary to reason strategically about likely actions of other players in games. Despite this obvious link, there is only a modest number of studies searching for activity in areas thought to be part of TOM in strategic games.

The first example is McCabe *et al.* (2001). They studied two-player trust games[8]. In a typical trust game, one player can end the game (giving both players 10, for example), or can trust a second player. Trust creates a larger collective gain (40), but the second player can share it equally (both get 20) or can keep it all (see Chapter 5 of this volume). Contrasting behavior with human partners and that with computer partners, they found that high-trust players had more activity in the paracingulate cortex and speculated that trust requires careful consideration of likely behavior of other players. Activity in the same general region is reported by Gallagher *et al.* (2002) in a PET "rock, paper, scissors" game when playing an experimenter rather than a computer opponent.

Given the link between autism and TOM, it is natural to use games to ask whether autists play differently

---

[8] Trust games are a sequential form of the well-known prisoners' dilemma (PD), with the modification that a defection by the first player always creates a defection by the second player. The sequentiality allows separation of trustingness and trustworthiness, which are confounded in the PD.

from control players. In the widely-researched ultimatum game, one player offers a share of money to another player, who can accept it or reject it. In these games, players typically offer 30–50% of the money, and offers that are too low are often rejected. Sally and Hill (2006) found that autists are much more likely to offer zero, apparently neglecting or misjudging the second player's move. Importantly, autistic children who offer positive amounts make a wide variety of offers, while positive offers by autistic adults consolidate around an equal split (similar to typical offers by normal adults). This consolidation of offers in adulthood around "normal" behavior is consistent with many reports that adult autists cope by learning explicit strategies for socially appropriate behavior.

Thinking Steps and Iterated Beliefs

The TOM evidence suggests that people are doing *some* strategic thinking, since playing humans versus computers activates TOM areas. The question raised by CH models, and their empirical success in explaining experimental data, is how *much* strategic thinking players do, what neural areas implement more strategic thinking, and related questions that arise.

Bhatt and Camerer (2005) compared player A's choices in games, A's expressed belief about B's choices, and A's "second-order" belief about B's belief about A's choice. Second-order beliefs are important in maintaining deception, because a successful deception requires A to make a certain choice and simultaneously believe that B believes he (A) will make a different choice. Second-order beliefs are also important in models of social image, in which a players' beliefs about what another player believes about his intentions or moral "type" influence utility (that is, players like to believe others believe they are good)[9].

One finding from Bhatt and Camerer's study is that in games in which a players' beliefs are all in equilibrium, there is little difference in neural activity when the player is making a strategy choice and expressing a belief. This suggests that the mathematical state of equilibrium (a correspondence of one player's beliefs with another player's actual choices) is also manifested by a "state of mind" – an overlap in the brain regions involved in choosing and guessing what others choose. They also find that second-order beliefs tend to err on the side of predicting that other players know what you will do, better than they actually do[10]. That is, players who planned to choose strategy S guessed that other players thought they would play S more often than the other players actually did. There is differential activity during the second-order belief task and first-order beliefs task in the insula, which has been implicated in the sensation of "agency" and self-causation and may help account for the self-referential bias in second-order beliefs.

So far, only one fMRI study has looked directly for neural correlates of the steps of thinking posited by the cognitive hierarchy model described earlier in the chapter. Coricelli and Nagel (2007) used a series of "beauty contest" number-choosing games (as shown in Figure 13.1). Players chose numbers 0–100 and have a target equal to p times the average number (for various p). Playing human opponents versus computers showed differential activity in medial paracingulate cortex and bilateral STS, as in other TOM studies. They classified players, using their choices, into low strategic reasoners (one step of reasoning, choices around p*50) and high strategic reasoners (two steps of reasoning, choosing around p^2*50). The high-step reasoners showed very strong differential activity (playing humans versus computers) in paracingulate, medial OFC, and bilateral STS (see Figure 13.3).

Since game theory presents many different tools to tap different aspects of strategic thinking, many more studies of this type could be done. The eventual goal is a mapping between the types of strategic thinking involved in games, components of theory of mind, and an understanding of neural circuitry specialized to each type of thinking.

Learning

Studies of the neuroscientific basis of learning in games fall into two categories. One category consists of attempts to see whether behavior exhibits some of the properties of reinforcement learning.

Seo and Lee (2007) recorded from monkey neurons in dorsal anterior cingulate cortex (ACC) during a matching-pennies game (the work–shirk game with equal payoffs in all cells) played against various computer algorithms. They found neurons with firing rates that are sensitive to reward and to some higher-order interactions with past choices and rewards. Behaviorally, the monkeys also play a little closer to the mixed equilibrium when the computer algorithms

---

[9] See Dufwenberg and Gneezy, 2002; Andreoni and Bernheim, 2007; Dillenberger and Sadowski, 2007; Ellingsen and Johannesson (2007) note the implications of this view for worker motivation in firms.

[10] This bias is related to psychological research on the "curse of knowledge" – the tendency of experts to think that novices know what they know (see any computer manual for evidence, or Camerer *et al.*, 1989) and the "illusion of transparency;" Gilovich *et al.*, 1998.

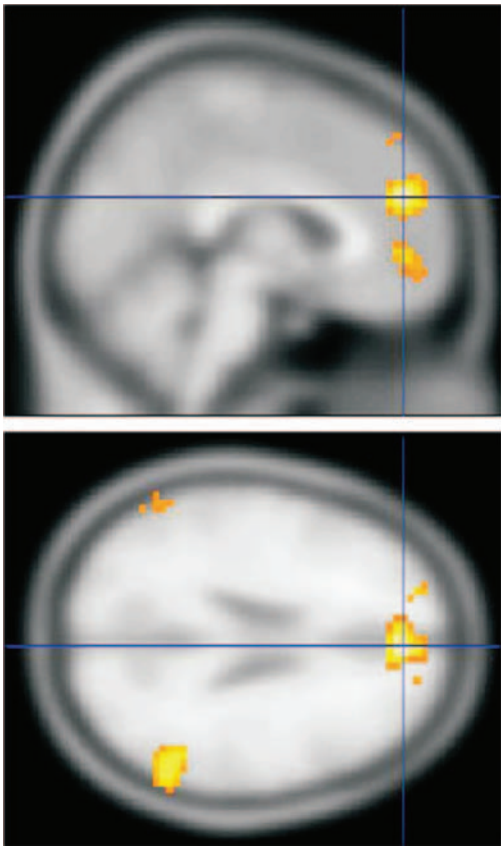


FIGURE 13.3 Differential activity in playing human vs computer opponents in "p-beauty contest" game for high-reasoning players. Players choose numbers 0–100 and winner earns p*(average number) for various values of *p*. High reasoners are those exhibiting "level 2" thinking (choosing closer to p^2 times 50 than p times 50). Enhanced activity for these subjects for human versus computer opponents is in paracingulate and medial paracingulate PFC (top), and bilateral STS (bottom). Reproduced from Coricelli and Nagel (2007), with permission.

are designed to exploit temporal dependence in the monkeys' play. Using the work–shirk game in Table 13.1, Dorris and Glimcher (2004) also found that monkeys play close to the mixed-equilibrium proportions, and adjust their strategy mixtures surprisingly rapidly, within 10–20 trials, when the game payoff parameters change. However, they note that neural firing rates in lateral intraparietal sulcus (LIP) do not change when strategies change, as long as the relative expected utility of strategies is the same. The LIP neurons are clearly encoding relative value, not choice rates.

The second category of neuroscientific studies explores generalizations of reinforcement which posit that learning can be driven by forces other than simply immediate reward. Lohrenz *et al.* (2007) define "fictive learning" as learning from counterfactual or imagined rewards (the d term in the EWA model). In an investment game (based on actual stock market prices), they show that a fictive learning signal is evident in caudate, close to a caudate region that encodes prediction error (the difference between outcome and expectation). The fictive signal also predicts changes in investment behavior.

Another interesting kind of learning arises when players engage in a repeated game. King-Casas *et al.* (2005) studied a repeated trust game. In the one-shot game, an "investor" player can invest an amount X from a stake of 20 which triples in value to 3X. The second, "trustee," player repays an amount Y, so the investor earns $(20 - X) + Y$ and the trustee earns $(3X - Y)$. Notice that the total payment is $20 + 2X$, so the collective payoff is maximized if everything is invested … but the investor cannot count on the trustee repaying anything. (Economists call this a game of investment with "moral hazard" and no enforcement of contracts, like investing in a country with poor legal protection.) King-Casas repeated the game 10 times with a fixed pair of players to study dynamics and learning, and scanned both investor and trustee brains simultaneously. Trustees tend to exhibit two kinds of behavior – they either reciprocate an uptick in investment from period $t - 1$ to t by repaying a larger percentage ("benevolent"), or reciprocate an uptick by investing less ("malevolent"). Figure 13.4 shows regions which are activated in the trustee choice period t by (later) benevolent trustee actions. The interesting finding, from a learning point of view, is that anticipation of benevolent "intention to trust" moves up by about 14 seconds from early rounds of the 10-period game (rounds 3–4) to later rounds 7–8. There is also both a within-brain correlation of this signal (trustee anterior cingulate and caudate) in anticipation of the later choice, and a cross-brain correlation (investor middle cingulate MCC and trustee caudate). That is, just as trustees are anticipating their own later benevolent action reward value, investors are anticipating it as well in medial cingulate cortex. This is a dramatic sign of synchronized anticipation due to learning, which could only be seen clearly by scanning both brains at the same time.

Hampton *et al.* (2007) used the work–shirk game (Table 13.1) to investigate neural correlates of "sophisticated" learning. Suppose an employer player, for example, has some inkling that others are learning from their own (employer) choices. Then if the employer chooses Inspect in one period, that choice has an immediate expected payoff (based on the employer's beliefs about what the worker will do) and also has a *future* influence because it is predicted to change the worker's beliefs and hence to change the worker's future play. Hampton *et al.* include this "influence value" as a regressor and correlate its numerical value with activity

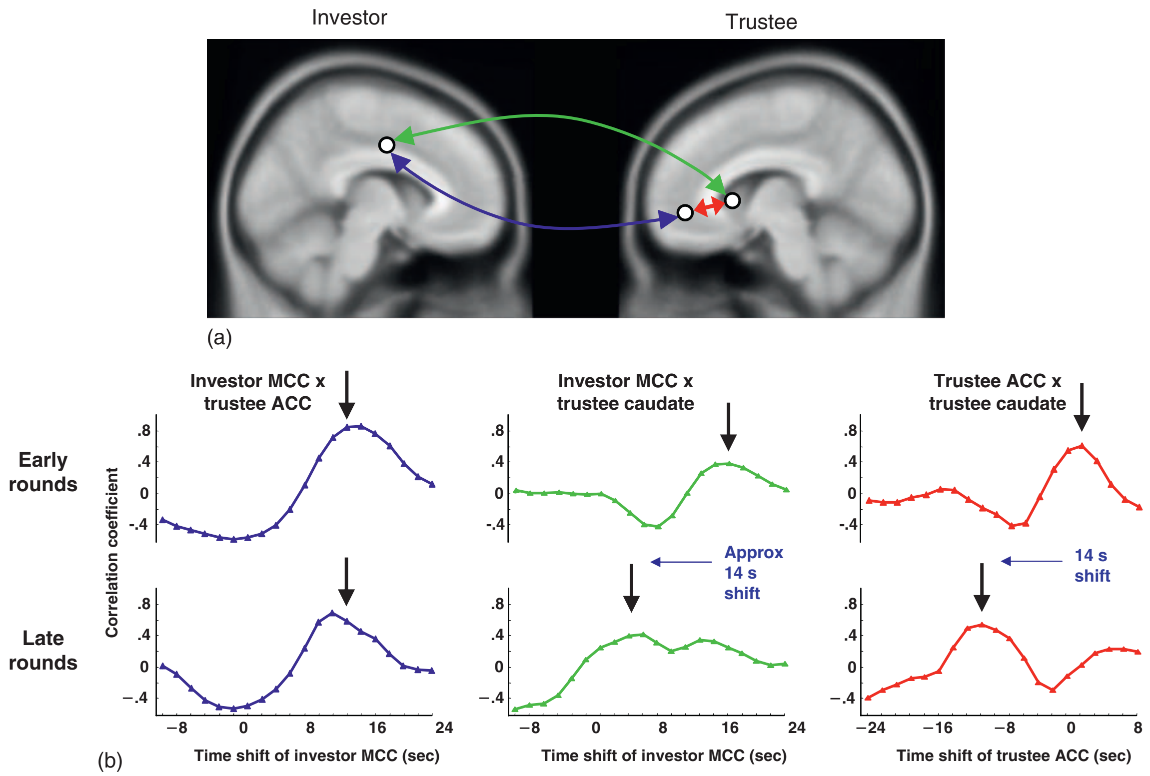


FIGURE 13.4 Trust-game activity in investor (first-moving player) and trustee (second-moving player) brains in 10-period trust game. (a) Regions activated in period *t* by "intention to trust" behavioral signal (reciprocal response to activity in the previous period) during trustee repayment phase. (b) Graphs show a time series of correlation between brain signals at different points in time. Positive correlations indicate two types of brain activity are highly correlated at the point in time indicated on the *y*-axis (0 is choice onset). Top time series is early rounds 3–4; bottom time series is later rounds 7–8. Correlations shift forward in time (~14 s) from early to late rounds, for both the cross-brain correlation of investor middle cingulate (MCC) and trustee caudate (left graphs), and within brain correlation of trustee ACC and trustee caudate (right graphs). The forward shift indicates that learning creates anticipation of likely behavior within the trustee's own brain (right graph) and between the two players' brains (left graph). Reproduced from King-Casas *et al*. (2005), with permission.

in the brain. The influence value (teaching) component activates posterior STS (Figure 13.5a) on a trial-by-trial basis. Furthermore, subjects can be categorized, purely from their behavioral choices, by how much better the influence model fits their choices than does a purely adaptive fictitious play model. There is a strong cross-subject correlation between the improvement in predicting behavior from including influence (Figure 13.5b *x*-axis) and activity in medial paracingulate in response to trial-by-trial influence value (Figure 13.5b *y*-axis). Along with the behavioral and eyetracking evidence, this finding provides direct fMRI evidence that human learning in games sometimes includes some degree of sophistication (as proposed and shown in behavioral data by Camerer *et al*., 2002, and Stahl, 2003).

Deception

Deception is an important topic that the combination of game theory and neuroscience may help illuminate. Game theory offers a rich set of games which characterize when deception is expected to occur, and potential field applications of these games. A useful aspect of the game theory is that it considers jointly the actions of a deceptive player and a player who can anticipate deception.

A useful mathematical example is "strategic information transmission" between an informed "sender" and a "receiver" (Crawford and Sobel, 1982). Consider a security analyst who has good information about the value V of a company, and a client who wants to know what V is. The analyst sends a message M to the client, who then chooses an action number A (a perception of V which triggers an investment policy). Often there is a conflict of interest in which the analyst wants the client to choose a number which is the true value plus a positive bias B. That is, the analyst's payoffs are highest if the client chooses an action A = V + B, and the client's payoff is highest if A = V. Notice that the message M is "cheaptalk:" in game theory language, this means that the message is costless (so the willingness

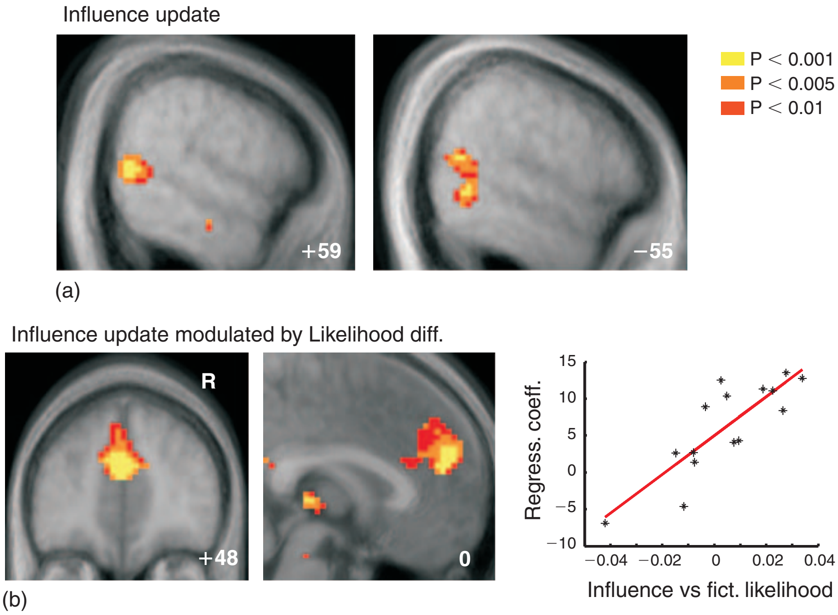


FIGURE 13.5 Correlation of the "influence value" of current action on future reward (through the influence on an opponent's future choice) with BOLD signal in fMRI. (a) Numerical influence values correlate with activity in bilateral STS trial-by-trial. (b) Cross-subject correlation of the extent to which choices are fit better by including an influence value term (x-axis) and strength of the influence value regressor in paracingulate cortex (y-axis). Reproduced from Hampton *et al*. (2007), with permission.

to pay for a message is not a calculated strategy that is informative) and does not bind behavior (so if the sender promises to do something and does not, there is no explicit penalty). The message number M can only influence payoffs through any information the client infers from the message which changes his action. This model has broad application to corporate advertising, expert services (fixing cars to fixing human bodies), political and personal promises, and bargaining.

When there is a large bias B, it can be proved mathematically that little truthful information is transmitted through the message M, if players are assumed to be self-interested and strategically sophisticated[11]. These equilibria are inefficient because if players could somehow agree to be truthful – or if a third party could certify information or penalize the analyst for exaggerating – then the players together would have a higher collective payoff.

Despite this strong prediction, several experiments have shown that in games with this structure, a substantial amount of information is communicated truthfully (and believed). There are two possible sources of this "overcommunication." One source is a feeling of guilt the sender has from exaggerating the truth and sending a message M which is higher than the value of V. Adding a negative utility from guilt into the analyst's payoffs can lead to an equilibrium with more truthful revelation. A second source is cognitive difficulty[12] it may just be hard to figure out how much to exaggerate, or how another player will react, so that heuristics like telling the truth or sending the message V + B (and hoping the client accepts it) are chosen. Wang *et al*. (2006) explored these causes in a sender-receiver game using eyetracking and pupillometry. They found that analysts do not look at the client's payoffs very often (although the equilibrium analysis predicts they have to, in order to figure out how receivers will react to different messages). Those who look at the client payoffs more often are not less deceptive, as a simple guilt theory predicts. Furthermore, the looking patterns and choices are mostly consistent with doing one or two steps of strategic thinking. Interestingly, using a combination of looking at information and pupil dilation when the analyst makes his choice gives a statistical prediction of the true state V which is sufficiently accurate to improve the analyst's experimental profits (on paper) by 10–20%.

Bhatt *et al*. (2007) have studied a closely-related game they call "yard sale bargaining," using fMRI.

---

[11] To illustrate, suppose the values are integers 1–5 and the bias is B = 1. Little information is transmitted, because any message about V that is truthful and believed will also be sent if the analyst has a lower value than V. For example, suppose when V = 3 the analyst truthfully announces M = 3, and the client believes it and chooses A = 3. Then an analyst who knows V = 2 will send the same message of 3, since he wants the client to choose V + B, which is 2 + 1 = 3. As a result, the equilibrium which conveys the most information is for the analyst to say M = 1 when V = 1 (admitting bad news) and to mix across 2–5 otherwise. When the bias is larger, B = 2, there is no information conveyed at all in equilibrium.

[12] The philosopher Friedrich Nietszche (1878/1996, p. 54) wrote "Why do men usually tell the truth in daily life? Certainly not because a god has forbidden lying. Rather it is because, first, it is more convenient: for lies demand imagination, dissembling, and memory (which is why Swift says that the man who tells a lie seldom perceives the heavy burden he is assuming: namely, he must invent twenty other lies to make good the first."

In this game, a seller has an item with value zero (so he would accept any positive price for it). A buyer has a value for the item V from 1–10, which the buyer knows but the seller does not. The buyer learns his value and suggests a price (a kind of cheaptalk, as in the analyst–client game), S. The seller sees S and then states a final take-it-or-leave-it price P. If the buyer's value is above the price $V \geq P$, the object is sold at the price P. If players are self-interested and strategic, there is no suggested price function S(V) which conveys any information. The reason is that any suggestion S which is believed to credibly convey the idea that the value is V would also be used by savvy buyers with values higher than V. So, in theory, the seller should completely ignore the suggestion S and state a price of 5–6 (which maximizes expected profits knowing nothing about V).

As in the analyst–client game, Bhatt *et al*. see that there is substantial revelation of information about value (contrary to the theory). A typical suggestion function is $S = V/2$ and a typical pricing function is $P = S + 2$. That is, the buyers often say they can pay half of what they can actually afford, and the sellers seem to guess this and pick a price which is the suggested price marked up by 2 units[13]. The fMRI shows that buyers who are less truthful about their values have greater activity in bilateral dorsal striatum (as if they were expecting larger profits). Sellers who are more sensitive to suggested price have greater activity in right anterior temporal sulcus, and less activity in the anterior cingulate. These regions are consistent with the hypothesis that believing messages is the default mode for sellers: Since the anterior cingulate is often involved in response conflict, lowered activity means the sellers who respond to suggested price are not registering a conflict between the suggested price and the likely buyer value. Responding to suggested prices activates the TOM temporal sulcus area (trying to infer the buyer's intention or state of mind), and ignoring those suggestions recruits ACC in order to resolve cognitive conflict (Kerns *et al*., 2004).

Reputation

An important concept in repeated game theory with private information is *reputation*. Private information is usefully characterized by a *type* a player has, which is randomly determined at the start of a repeated game. In theory, a player's actions are designed to satisfy short-term goals and also to either convey (in a cooperative) game or hide (in a competitive game) the player's type. Player A's reputation is the belief, in the eyes of other players, about player A's type.

For example, an important analysis of the repeated prisoners' dilemma (PD) starts with the idea that some players always cooperate. Then, even players who are selfish will choose to cooperate, in order to maintain a reputation as the kind of player who cooperates, because having such a reputation encourages cooperation from other players in the future.

Two neural studies indirectly tap aspects of reputation. Singer *et al*. (2004) found that showing faces of people who had previously cooperated activated the nucleus accumbens. This is the first direct evidence that a game-theoretic reputation generates a neural value signal. Delgado *et al*. (2005) used fMRI to explore neural reactions to behavior in a repeated cooperation game when the scanned subject's opponent begins with a good, neutral, or bad reputation created by a picture and short blurb about an opponent's behavior (in Bayesian terms, the blurb creates a prior belief that the opponent will behave cooperatively or not). They found that during the outcome phase, if the partner behaves cooperatively, compared to uncooperatively, there is differential activity in the caudate nucleus (and several other areas). However, there is no such difference in this contrast if the partner had a good reputation to begin with. The time course of activity is consistent with the idea that bad behavior is "forgiven" (in neural terms, does not generate as much reward or prediction error signal) if the partner is a good person.

CONCLUSIONS AND FUTURE RESEARCH

Game theory is useful for creating a precise mathematical model linking strategy combinations to payoffs, a kind of periodic table of the elements of social life. Predictions are made using various behavioral assumptions about how deeply people reason and how they react to observed behavior. Hundreds of experiments suggest that players do not always reason very strategically, evaluation of payoffs often includes social elements beyond pure self-interest, and players learn from experience.

So far, there has been only limited use of game theory and neuroscientific tools to link strategic thinking to neural activity. This limited contact is probably due to the fact that psychologists have not used the major tools in game theory, which may in turn be due to skepticism that the rationality-based analyses in game theory are psychologically accurate.

---

[13] Mathematically, these strategies imply that trade takes place when $V \geq P$, or $V \geq (V/2) + 2$ which implies $V \geq 4$, so that more trades take place than would in equilibrium.

One promising point of contact is between theories of strategic thinking and "theory of mind" (TOM) regions thought to be necessary for understanding beliefs, desires, and thoughts of other people. The few available studies tend to indicate that TOM areas are activated in playing mathematical games, but a closer link would be very useful for both fields.

Game theory could also be useful in understanding disorders. Some psychiatric disorders could be understood as disorders in normal social evaluation and prediction. For example, anti-social personality disorder seems to disrupt normal valuation of the consequences of one's actions on others. Paranoia in psychosis and schizophrenia could be defined symptomatically as overpredicting a hostile (payoff-reducing) reaction of others to one's own choices. Autism can also be seen as a disorder in evaluating expected social behavior. Using a battery of games involving altruism, fair sharing, and trust, Krajbich *et al*. (2008) have found that patients with ventromedial prefrontal cortical damage act as if they exhibit less parametric guilt – giving less and acting in a less trustworthy fashion – than do normal controls and control patients with damage in other regions.

Game theory is also a tool for understanding expertise and increasing skill. In a game, there is usually a clear performance metric – who makes the most money? Understanding extraordinary skill in bargaining, poker, and diplomacy may illuminate the everyday neural bases of these skills and permit effective training.

References

Aumann, R. (1985). What is game theory trying to accomplish? In: K. Arrow and S. Honkaphoja (eds), *Frontiers of Economics*. Oxford: Basil Blackwell, pp. 28–76.

Andreoni, J., and Bernheim, B.D. (2007). Social image and the 50–50 norm: a theoretical and experimental analysis of audience effects. August, http://econ.ucsd.edu/~jandreon/WorkingPapers/socialimage.pdf

Behrens, T.E.J., Woolrich, M.W., Walton, M.E., and Rushworth, M.F.S. (2007). Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221.

Bhatt, M. and Camerer, C.F. (2005). Self-referential thinking and equilibrium as states of mind in games: fMRI evidence. *Games Econ. Behav.* 52, 424–459.

Bhatt, M., Lohrenz, T., Montague, R.M., and Camerer, C.F. (2007). Neural correlates of lowballing and gullibility in "yard-sale bargaining". Working Paper, Caltech.

Brown, A.L., Camerer, C.F., and Lovallo, D. (2007). *To review or not review? Limited strategic thinking at the box office*. Pasadena, CA: California Institute of Technology.

Camerer, C.F. (1998). *Mental Representations of Games*. Princeton, NJ: Princeton University Press.

Camerer, C.F. (2003). *Behavioral Game Theory: Experiments on Strategic Interaction*. Princeton, NJ: Princeton University Press.

Camerer, C.F. and Fehr, E. (2006). When does "Economic Man" dominate social behavior? *Science* 311, 47–52.

Camerer, C. and Ho, T.H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica* 67, 827–874.

Camerer, C., Loewenstein, G., and Weber, M. (1989). The curse of knowledge in economic settings – an experimental analysis. *J. Political Econ.* 97, 1232–1254.

Camerer, C.F., Johnson, E., Rymon, T., and Sen, S. (1993). Cognition and framing in sequential bargaining for gains and losses. In: K.G. Binmore, A.P. Kirman, and P. Tani (eds), *Frontiers of Game Theory*. Cambridge: MIT Press, pp. 27–47.

Camerer, C.F., Ho, T.-H., and Chong, J.-K. (2004). A cognitive hierarchy model of games. *Q. J. Economics* 119, 861–898.

Camerer, C.F., Rogers, B., and Palfrey, T. (2008). Heterogeneous quantal response equilibrium and cognitive hierarchies. *J. Econ. Theory*, (in press).

Chong, J.-K., Camerer, C., and Ho, T.-H. (2006). A learning-based model of repeated games with incomplete information. *Games Econ. Behav.* 55, 340–371.

Coricelli, G. and Nagel, R. (2007). *Guessing in the Brain: An fMRI Study of Depth of Reasoning*. Working Paper, Lyon University.

Costa-Gomes, M.A. and Crawford, V.P. (2006). *Cognition and Behavior in Two-Person Guessing Games: An Experimental Study*. London: UCLA, Department of Economics.

Costa-Gomes, M.A., Crawford, V.P., and Broseta, B. (2001). Cognition and behavior in normal-form games: an experimental study. *Econometrica* 69, 1193–1235.

Crawford, V.P. and Sobel, J. (1982). Strategic information transmission. *Econometrica* 50, 1431–1451.

Delgado, M.R., Frank, R.H., and Phelps, E.A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8, 1611–1618.

Devetag, G. and Warglien, M. (2003). Games and phone numbers: do short-term memory bounds affect strategic behavior? *J. Econ. Psychol.* 24, 189–202.

Dillenberger, D. and Sadowski, P. (2007). *Ashamed to Be Selfish*. Princeton, NJ: Princeton University.

Dorris, M.C. and Glimcher, P.W. (2004). Activity in posterior parietal cortex is correlated with the subjective desirability of an action. *Neuron* 44, 365–378.

Dufwenberg, M. and Gneezy, U. (2000). Measuring beliefs in an experimental lost wallet game. *Games Econ. Behav.* 30, 163–182.

Ellingsen, T. and Johannesson, M. (2007). Paying respect. *J. Econ. Persp.* 21, 135–149.

Fehr, E. and Camerer, C.F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* 11, 419–427.

Frith, C.D. and Frith, U. (2006). The neural basis of mentalizing. *Neuron* 50, 531–534.

Gallagher, H.L., Jack, A.I., Poepstorff, A., and Frith, C.D. (2002). Imaging the intentional stance in a competitive game. *NeuroImage* 16, 814–821.

Gilovich, T., Savitsky, K., and Medvec, V.H. (1998). The illusion of transparency: giased assessments of others' ability to read our emotional states. *J. Pers. Social Psychol.* 75, 332–346.

Goeree, J.K. and Holt, C.A. (2001). Ten little treasures of game theory and ten intuitive contradictions. *Am. Econ. Rev.* 91, 1402–1422.

Hampton, A., Bossaerts, P., and O'Doherty, J. (2007). Neural correlates of mentalizing-related computations during strategic interactions in humans. Working Paper, Caltech.

Hedden, T. and Zhang, J. (2002). What do you think I think you think? Strategic reasoning in matrix games. *Cognition* 85, 1–36.

Ho, T.H., Camerer, C.F., and Weigelt, K. (1998). Iterated dominance and iterated best response in experimental "p-beauty contests". *Am. Econ. Rev.* 88, 947–969.

Ho, T.H., Camerer, C.F., and Chong, J.-K. (2007). Self-tuning experience weighted attraction learning in games. *J. Econ. Theory* 127, 177–198.

Johnson, E.J., Camerer, C., Sen, S., and Rymon, T. (2002). Detecting failures of backward induction: monitoring information search in sequential bargaining. *J. Econ. Theory* 104, 16–47.

Kerns, J.G., Cohen, J.D., MacDonald, A.W., III *et al*. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science* 303, 1023–1026.

King-Casas, B., Tomlin, D., Anen, C. *et al*. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83.

Krajbich, I., Adolphs, R., Tranel, D. *et al*. (2008). Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. Working Paper, Caltech.

Lohrenz, T., McCabe, K., Camerer, C.F., and Montague, P.R. (2007). Neural signature of fictive learning signals in a sequential investment task. *PNAS* 104, 9493–9498.

McCabe, K., Houser, D., Ryan, L. *et al*., (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl Acad. Sci. USA*, 98, 11832–11835.

McKelvey, R.D. and Palfrey, T.R. (1995). Quantal response equilibria for normal form games. *Games Econ. Behav.* 10, 6–38.

McKelvey, R.D. and Palfrey, T.R. (1998). Quantal Response equilibria for extensive form games. *Exp. Economics* 1, 9–41.

Myerson, R.B. (1986). Acceptable and predominant correlated equilibria. *Intl J. Game Theory* 15, 133–154.

Nagel, R. (1995). Unraveling in guessing games: an experimental study. *Am. Econ. Rev.* 85, 1313–1326.

Nietzsche, F. (1996). *Human, All Too Human: A Book for Free Spirits*. Cambridge: Cambridge University Press.

Östling, R., Wang, J.T.-y., Chou, E., and Camerer, C.F. (2007). Field and lab convergence in Poisson LUPI games. Working Paper Series in Economics and Finance. Stockholm: Stockholm School of Economics.

Sally, D. and Hill, E. (2006). The development of interpersonal strategy: autism, theory-of-mind, cooperation and fairness. *J. Econ. Psychol.* 27, 73–97.

Saxe, R. and Powell, L.J. (2006). It's the thought that counts. *Psychological Sci.* 17, 692–699.

Schweighofer, N. and Doya, K. (2003). Meta-learning in reinforcement learning. *Neural Networks* 16, 5–9.

Seo, H. and Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* 27, 8366–8377.

Singer, T., Kiebel, S.J., Winston, J.S. *et al*. (2004). Brain responses to the acquired moral status of faces. *Neuron* 41, 653–662.

Soltani, A., Lee, D., and Wang, X.-J. (2006). Neural mechanism for stochastic behaviour during a competitive game. *Neural Networks* 19, 1075–1090.

Stahl, D.O. (2003). Sophisticated learning and learning sophistication. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=410921

Stahl, D.O. and Wilson, P.W. (1995). On players' models of other players: theory and experimental evidence. *Games Econ. Behav.* 10, 218–254.

Wang, J.T.-y., Spezio, M., and Camerer, C.F. (2006). Pinocchio's pupil: using eyetracking and pupil dilation to understand truth-telling and deception in biased transmission games. Pasadena, CA: Caltech.

Wang, J. T.-y., Knoepfle, D., and Camerer, C.F. (2007). Using eye-tracking data to test models of learning in games. Working Paper, Caltech.

PART III

SOCIAL DECISION MAKING, NEUROECONOMICS, AND EMOTION

C H A P T E R

14

Neuroscience and the Emergence of Neuroeconomics

Antonio Damasio

O U T L I N E

It is certainly the case that in the mid-1990s the term *neuroeconomics* was not in use, and that the field of studies that now goes by that name did not yet exist. However, the foundational facts were available, the key ideas were in the air, and all were ready to be focused on a new target. It takes two to dance, and in this case there were indeed two partners: behavioral economics and neuroscience.

As I see it, the behavioral economics partner was by then well established and had contributed a central idea – namely, that rational choice could not account satisfactorily for a considerable number of economic behaviors. To substantiate this idea, behavioral economics had gathered a remarkable roster of facts (for review, see Kahneman, 2003).

The neuroscience partner had also contributed a combination of facts and ideas. I will review some of those, from today's perspective, and round up my comments with a reflection on the notion of biological value, an indispensable construct in neuroeconomics.

FROM NEUROSCIENCE TO NEUROECONOMICS

A number of neuroscience developments were of special relevance in the emergence of the new field of neuroeconomics, and I will begin by highlighting those that have to do with the neural basis of decision making. A brief review of the critical evidence reveals the following facts which came to light between the mid-1980s and the mid-1990s:

1. Previously normal individuals who sustained bilateral brain damage centered on the ventral and medial sectors of the prefrontal cortices exhibited, after the onset of damage, marked defects of decision making. The defects were especially notable for social behaviors.
2. In two areas of social behavior, the defects were so evident that they practically required no special diagnostic tool; these areas were *interpersonal*

relationships and, notably, decision making having to do with *financial* issues. (Curiously, these findings already foreshadowed the main bodies of research that were to spring from them: social neuroscience and its exploration of moral aspects of behavior (examples of which can be found in Tania Singer's work and in the work of my research group); and the neuroscience of economic behaviors, a specialization of the pursuit into the neural underpinnings of social behavior, of which Ernst Fehrs' work is a great example.

3. The patients with ventromedial prefrontal lesions had remarkably preserved intellect, as measured by conventional neuropsychological instruments, and an equally remarkable defect of emotional behavior. The emotional defect consisted of a rather general diminished emotional resonance, along with specific and notable impairments in social emotions – for example, in compassion and embarrassment.

In brief, patients, who had had normal social behavior until the onset of their brain dysfunction, and who certainly had not had any comparable difficulties in making sound decisions until lesion onset, were now deciding poorly and generally doing so against their best interests and the interests of those closest to them. This was happening in spite of their intellectual instruments being essentially preserved. The patients had no detectable impairments of logical reasoning, no defects of learning and recall of the kind of knowledge required to make sound decisions, and no defects of language or perception. Yet their decisions were flawed. Upon having the flaw pointed out to them, they did recognize that they could have done better. Once placed in similar future situations, however, they were likely to make comparably defective decisions.

The contrast between defective emotion on the one hand and preserved intellect on the other led me to propose that, somehow, disturbed emotional signaling could explain the decision defect. This idea formed the basis for the so-called *somatic marker hypothesis*, which was aired in several articles during the 1990s. Easily accessible summaries of findings and theory can be found in *Descartes' Error* (Damasio, 1994), and in an article for the *Transactions of the Royal Society* (Damasio, 1996).

I never considered the hypothesis as anything but a beginning, the start of an exploration of the role of emotion in decision making, but I did think that such a possibility was well worth entertaining; namely, that emotion would play an important role in decision making not just for the worst, as was then the traditional view, but for the *better*. I was persuaded that emotion might well account for some of the decision anomalies brought to light by the work of Kahneman and Tversky. And I did note, from the outset, that the "emotion" concept used in the theory was nothing but the tip of the iceberg. Underneath that iceberg there were the mechanisms of drives and motivations as well as those of reward and punishment, which are the fundamental constituents of the emotion machinery. I ventured that those were the factors most likely to play the main modifying role in the decision process, from a neural perspective, at either conscious or unconscious level.

In retrospect, it is apparent that these early observations and interpretations benefited from and became part of a major revival of interest in the neuroscience of the emotions, which had been much neglected until the last decade of the twentieth century. The new work on the emotions encompassed research in experimental animals – a prime example is Joseph Le Doux's exploration of the fear conditioning paradigm in rodents (Le Doux, 1996) – as well as the human lesion studies conducted by our group. Over a brief period of time a growing number of investigators were able to identify critical stages in the emotional process, and discover the main induction sites for emotions such as fear (the amygdaloid nuclei) and the social emotions (the ventromedial prefrontal cortices). We were also able to establish a principled distinction between emotion and feeling (see below), and to identify the insular cortex as a principal neural substrate for feelings (Damasio, 1994; Damasio *et al*., 2000). Social neuroscience and neuroeconomics were by then ready to exploit functional neuroimaging to its full advantage, a trend that has continued unabated.

Adopting today's neuroeconomics perspective, I would summarize the somatic-marker hypothesis as follows:

1. Emotion plays a role in decision making, but it should be clear that, under the term *emotion* I include both (a) the neural subprocesses of automated life regulation that are part and parcel of emotion action programs, namely reward and punishment processes and drives and motivations; and (b) the neural substrates of the perceptual read-outs of emotion action programs, namely emotional *feelings*.
2. In the original somatic-marker hypothesis outline, I suggested that the emotional influence on the decision-making process was exerted neurally, at multiple neural levels, from the high level of feelings substrates to the level of reward and punishment signaling (see Damasio, 1996). Needless to say, I remain convinced of the importance of these points and I wish to emphasize them because

so often, especially in discussions on the notion of biological value, the concept of emotion becomes dangerously amputated. Separating emotion from its reward and punishment components is a major conceptual problem. Another major conceptual problem comes from confusing emotion (which is an action program) with a *feeling* of emotion (which is the conscious, cognitive sequel to the action program). These are different phenomena with different neural substrates.

3. Emotion plays its role either consciously or non-consciously, depending on the stage of the process and the circumstances. When emotion influences decisions consciously, the deciding subject may be aware of the "marker" and even refer to it – for example, report a "gut feeling." But decisions may also be influenced covertly, and the hypothesis states that non-conscious "biases" can alter the processing networks and drive the process in a particular direction. I conceived of this, and still do, as operated by specific neuromodulators acting on different levels of neural circuitry, all the way to the cerebral cortex.
4. In the framework of the somatic marker hypothesis, the abnormal decision making that we described in our patients resulted from a cognitive malfunction that was rooted in an emotional malfunction. In other words, the emotional defect did not explain the anomaly alone; the emotional malfunction altered the cognitive process.
5. The term "somatic" needs some clarification. It conjured up the *body-relatedness* of the physiological mechanisms I was invoking. I believed, and still do, that the decision-making machinery we make use of in all social matters recruits, and in economic matters in particular, mechanisms of decision making that began as routines of life regulation focused on body physiology. Hence the word *somatic*.
6. What I meant by *marker* in the somatic-marker hypothesis is sometimes misinterpreted. The marker in the hypothesis is a *memory trace*. The marker was learned in past experiences of the subject, in which certain situations (a) required a decision, (b) evoked certain options of action, (c) prompted a decision, and (d) resulted in specific outcome. The outcome would have been, in the emotional sense, positive or negative, rewarding or punishing. In other words, the marker stands for situations in which certain facts (the premises of a problem; the options of action; the factual outcome) were associated with certain emotional outcomes. The marker signals the conjunction, in past experience, of certain categories of situation or outcome with certain categories of emotional response. The marker as memory trace is recorded in higher-order cortical circuitry, of which the ventro-medial prefrontal cortices are the most notable example.
7. When situations of a certain category re-present themselves to the decider subject, the marker is reactivated. In other words, processing a situation strongly resembling another situation regarding which decisions were made, prompts recall of related information. The recall may or may not come to consciousness, but in either case it promotes the replication, partial or complete of the emotional state associated with the particular class of situation, option, or outcome. In normal individuals, the marker "weighs in" on the decision process. In cases of ventromedial prefrontal damage, it fails to do so.

The somatic market hypothesis prompted several experimental tests of its validity, and inspired the development of the Gambling Task (Bechara *et al.*, 1994). The task provided the first laboratory diagnostic procedure for patients with ventromedial prefrontal damage – a rather useful advance, given that these patients generally passed all other neuropsychologic tests and only exhibited their defects in real life and real time. The task was also instrumental in showing a persuasive correlation between indices of emotional change (skin conductance responses) and the advantageous or disadvantageous playing of the card game (Bechara *et al.*, 1997). The poor performance of prefrontal patients was accompanied by largely flat skin conductance responses which failed to discriminate between advantageous and disadvantageous decks.

The task attracted an intriguing controversy regarding how conscious the normal individuals who played the card game were of the winning strategy. When critics Maia and McClelland (2004) administered the Gambling Task using our procedures they replicated our results, as has been the case with all other authors who have done so. However, when Maia and McClelland used a different set of instructions for the task, one that probed ongoing knowledge in a deeper manner, the results predictably revealed that the subjects knew about the winning strategy earlier than in our version. The deeper probing was responsible for a greater scrutiny of the task by the subject, and injected into the process a degree of knowledge that our version of the procedures did not prompt. In no way do the results of the modified task contradict our original task results, or the idea that deciders, in the Gambling Task or in other situations, may be influenced by non-conscious factors, emotional or otherwise (a recent study by Persaud *et al.*, 2007, bears out this point nicely). And in no way does the modified task compromise the somatic marker hypothesis, since the hypothesis specifies that the emotional role in decisions can be played out either consciously or non-consciously.

DECISION MAKING, EMOTION, AND BIOLOGICAL VALUE

I will conclude by turning to the issue of biological value. Neuroscience has identified several chemical molecules that are, in one way or another, associated with value – dopamine, cortisol, oxytocin, and prolactin. Neuroscience has also identified a number of subcortical neuron nuclei, located in the brainstem and hypothalamus, which manufacture those molecules and deliver them to selected parts of the brain and of the body. The complicated neural mechanics of those molecules is an important topic of neuroscience that many committed researchers wish to unravel. What prompts the release of those molecules? Where do they go exactly? What do they accomplish? But somehow, discussions about all the new facts come up short when turning to the most central questions:

1. Where is the engine for the value systems
2. What is the biological primitive value?

We need to know why things came to be this way.

The gist of my answers is as follows. Value is indelibly tied to need, and need is tied to life. The valuations we establish in everyday social and cultural activities have a direct or indirect connection with human biology and, in particular, with the processes of life regulation known by the term *homeostasis*. Value relates, directly or indirectly, to survival.

Because survival means different things in the perspectives of genes, cells, systems, whole organisms, and cultures, the origins of value will appear to be different depending on the target of the observation. Let me begin by considering the whole-organism level.

The machinery of homeostasis has been designed (obviously, by "designed" I mean achieved by selectional processes over evolutionary time) to protect the integrity of living organisms and, to state it crudely, as far as organisms go, the paramount value for these organisms consists of healthy survival to an age compatible with procreation. Accordingly, I regard the physiological state of tissues within a living organism, specifically, *the state of living tissue within a homeostatic range*, as the deepest origin of biological value and valuations. Toward one extreme of the homeostatic range the viability of living tissue declines and the risk of disease and death increases; toward the other extreme of the range, living tissue flourishes and its function becomes more efficient and economic. States closer to the former extreme are less valuable than states in the middle range and states closer to the latter extreme. The primitive of organism value is inscribed in the physiological parameters of the state itself. It is plausible that other processes and objects acquire their assigned value by reference to this primitive of organism value.

The values attributed to objects and activities will bear some relation, no matter how indirect or remote, to the maintenance of living tissue within a homeostatic range. As noted, the neurobiology literature tends to be vague regarding the issue of value. Some accounts mention the machinery of punishment and reward as the basis of value; some remind us of the chemical molecules related to such machinery; and most tend to overlook the fact that emotion needs to be part of the picture.

The origins of value as outlined for a whole organism apply quite well to an individual cell. Value is still defined in terms of a physiological state. However, it is reasonable to wonder how the conditions described for cells and organisms come to be. To approach such an issue, we must consider events that took place in a long ago evolutionary past – a reverse form of engineering that is never easy. Humans have spent most of their scientific history observing whole organisms and their major components while obfuscating, down below, the gene level where each organism began. And that is the level we must go to in order to discover where the power of homeostasis originates.

We can begin by considering that, in order to continue their existence over generations, gene networks needed to construct perishable, complex, and yet successful organisms that served as vehicles for their advancement; and that in order for organisms to behave in that successful manner genes must have guided the design of those organisms with some critical instructions.

My hypothesis is that a good part of those fundamental instructions ended up constructing devices capable of conducting general life regulation (homeostasis), distributing rewards, applying punishments, and helping predict the next situation of an organism – in brief, devices capable of executing what we have come to call emotions, in the broad sense of the term. The early sketch of these devices was first present in organisms without mind or consciousness, in fact without a brain, but the regulating devices attained the greatest complexity in organisms that do have all three: brain, mind, and consciousness. I suspect the empowering instructions were an important engine early in evolution, and everything suggests that they are still in use today, from the level of operations that regulates our metabolism to the level of human behaviors present in sociopolitical activities and, of course, in economics in the narrow sense of the term.

References

Bechara, A., Damasio, A.R., Damasio, H., and Anderson, S.W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 50, 7–15.

Bechara, A., Damasio, H., Tranel, D., and Damasio, A. (1997). Deciding advantageously before knowing the advantageous strategy. *Science* 275, 1293–1295.

Damasio, A.R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York, NY: Penguin Books.

Damasio, A.R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Phil. Trans. R. Soc. Lond. B* 351, 1413–1420.

Damasio, A.R., Grabowski, T.J., Bechara, A. *et al*. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nat. Neurosci.* 3, 1049–1056.

Kahneman, D. (2003). Maps of bounded rationality: psychology for behavioral economics. *Am. Econ. Rev.* 93, 1449–1475.

Le Doux, J. (1996). *The Emotional Brain*. New York, NY: Simon and Schuster.

Maia, T.V. and McClelland, J.L. (2004). A reexamination of the evidence for the somatic marker hypothesis: what participants really know in the Iowa gambling task. *Proc. Natl Acad. Sci. USA* 101(16), 075–16,080.

Persaud, N., McLeod, P., and Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nat. Neurosci.* 10, 257–261.

CHAPTER

15

Social Preferences and the Brain

Ernst Fehr

OUTLINE

INTRODUCTION

Many influential economists have pointed out that people not only care about their own welfare but also for the well-being of others, and that this may have important economic consequences; these include one of the founding fathers of economics, Adam Smith (Smith, 1976), and Nobel Prize winners such as Gary Becker (Becker, 1974), Kenneth Arrow (Arrow, 1981), Paul Samuelson (Samuelson, 1993), Amartya Sen (Sen, 1995) and Reinhard Selten (Selten, 1998). Throughout most of its history, however, mainstream economics has relied on the simplifying assumption that material self-interest is the *sole* motivation of *all* people, and terms such as "other-regarding preferences" were simply not part of economists' vocabulary. The term "social preferences" was occasionally used to discuss the problem of assigning a preference to an aggregate entity such as a whole group or society. The question here was a problem of aggregation – how to derive the preference of a whole group of people (the "social" preference) from the individual group members' preferences.

One reason for the prevalence of the self-interest hypothesis in economics is that it has served the profession quite well because self-interest is without doubt one important motivational force and some people indeed display very self-interested behaviors. In some domains, such as competitive experimental markets, models based on the self-interest hypothesis even make very accurate quantitative predictions (Smith, 1962, 1982). However, in strategic interactions, where individuals' actions typically have a direct impact on other individuals' payoffs, the self-interest hypothesis often fails to predict correctly (Fehr and Gächter, 2000; Camerer and Fehr, 2006). These predictive failures of the self-interest model gave rise to the development of social preference models (see Box 15.1).

A "social preference" is now considered to be a characteristic of an individual's behavior or motives, indicating that the individual cares positively or negatively about others' material payoff or well-being. Thus, a social preference means that the individual's motives are other-regarding – that is, the individual takes the welfare of other individuals into account.

BOX 15.1

FORMAL THEORIES OF SOCIAL PREFERENCES

Formalization brings rigor to science. Therefore, economists have developed formal models of social preferences that describe motivational forces precisely and transparently. Almost all models are based on a utility functions of the form $U_i = x_i + \sum_j v_i^j \cdot x_j$, where U_i is the utility of player i, x_i is the material payoff of player i, and the summation is over all $j \neq i$. The term v_i^j measures player i's valuation of player j's payoff. If v_i^j is negative, j's payoff is valued negatively so that i is willing to incur costs to reduce j's payoff, If v_i^j is positive, j's payoff is valued positively so that i is willing to incur costs to increase j's payoff. v_i^j is always zero for selfish players. Below are presented four important formalizations of social preferences, each of which highlights one aspect of other-regarding motives. Evidence in favor and against the different approaches is given in Fehr and Schmidt, (2003).

In theories of reciprocity, v_i^j depends on j's kindness to i, while it depends on the payoff difference between i and j in theories of inequity aversion. More formally, in the case of reciprocal preferences, U_i is given by $U_i = x_i + \sum_j v_i(\kappa_i^j) \cdot x_j$, and the term κ_i^j measures player j's kindness towards player i. In case of inequity averse preferences, κ_i^j is determined by the prevailing difference in material payoffs between i and j.

Menu-based reciprocity (Rabin, 1993). In menu-based models, j's kindness is determined by the actual choice of j in comparison to the alternatives (the available menus). Let A_i^j denote the set of available alternatives, which determine the possible payoffs available to player i depending on player j's choice. Let π_j^L be the lower payoff limit of A_i^j and π_i^H the upper limit of A_i^j. We define the fair payoff as $\pi_i^F = (\pi_i^H + \pi_i^L)/2$. Let π_i^A be the payoff of player i given the actual choice of player j. The kindness κ_i^j of player j toward i is defined as 0 if $\pi_i^H = \pi_i^L$ and as $2(\pi_i^A - \pi_i^F)/(\pi_i^H - \pi_i^L)$ otherwise. This expression is always between -1 and $+1$. The evaluation function in this model is simply the multiplication of κ_i^j with an individual reciprocity parameter $\rho_i \geq 0$, which measures the weight of the reciprocity motive. The utility of player i in the two-player case is therefore defined as $U_i = x_i + \rho_i \kappa_i^j x_j$ which is determined by the actions and the beliefs of the players. A reciprocity equilibrium is then defined as a combination of actions and beliefs in which, first, all players choose a strategy to maximize their utility and, second, beliefs match the actual behavior.

Outcome-based fairness. In this model, $\kappa_i^j = x_i - x_j$ and the evaluation function is given by

$$v_i(\kappa_i^j) = \begin{cases} +\beta_i/(n-1) & \textit{if} \quad \kappa_i^j > 0 \\ 0 & \textit{if} \quad \kappa_i^j = 0, \\ -\alpha_i/(n-1) & \textit{if} \quad \kappa_i^j < 0 \end{cases}$$

where n represents the number of players and $\alpha_i > 0, \beta_i > 0$ for a fair player. The above model mimics reciprocal fairness, i.e. j's payoff is valued positively if j is worse off, and negatively if j is better of than i. Based on this definition, outcome-based reciprocal fairness can be transformed into inequity aversion by assuming a utility function $U_i = x_i + \sum_j v_i(\kappa_i^j) \cdot (x_j - x_i)$, which is the function stipulated by Fehr and Schmidt (1999) if one imposes the parameter restrictions $\alpha_i \geq \beta_i \geq 0$ and $\beta_i < 1$. Note that v_i no longer weights the other player's payoff but inequality, and α_i measures the disutility from being worse off (envy) while β_i measures the disutility of being better off (compassion). In the two-player case, this utility function simplifies to $U_i = x_i - \alpha_i(x_j - x_i)$ if j is better off than i, and $U_i = x_i - \beta_i(x_i - x_j)$ if i is better off than j.

Personality-based reciprocity (Levine, 1998). Assume that players differ in how altruistic they are and that their degree of baseline altruism can be captured by the parameter α_i. Personality-based theories assume that people predict other individuals' altruism parameter. They respond with altruistic rewarding or altruistic

BOX 15.1 (*Cont'd*)

punishment, depending on their prediction of others' altruism parameters. More formally, the utility payoff of such players is given by

$$U_i = x_i + \sum_j \frac{\alpha_i + \lambda_i \cdot \alpha_j}{1 + \lambda_i} x_j$$

where α_i captures player *i*'s altruistic motivation (and obeys $-1 < \alpha_i < 1$) and the reciprocity parameter λ_i measures player *i*'s preference for reciprocation (and obeys $0 \leq \lambda_i \leq 1$). Here kindness κ_i^j is defined by player *j*'s altruism parameter α_j. The valuation function v_i is given by $v(\kappa_i^j) = (\alpha_i + \lambda_i \kappa_i^j)/(1 + \lambda_i) = (\alpha_i + \lambda_i \alpha_j)/(1 + \lambda_i)$. This model has two key properties. First, the higher α_i, the more player *i* values the other players' payoff. If $\alpha_i < 0$, player *i* is even spiteful, i.e., he prefers reducing the other player's economic payoff. Second, the higher the altruism parameter of the other player *j*, the more a reciprocal player *i* (with $\lambda_i > 0$) values player *j*'s economic payoff.

Rawlsian preferences and preferences for the group's overall payoff (Charness and Rabin, 2002). This approach combines preferences for the group's overall material welfare ("efficiency") with a Rawlsian version of inequity aversion (Rawls, 1972) in which a player cares only for the worst-off player's payoff. The utility function in this case is given by

$$U_i = x_i + \gamma[\delta \cdot min\{x_{1,} \ldots, x_n\} + (1 - \delta) \cdot \Sigma x_j].$$

where $\gamma > 0$ and $0 < \delta < 1$. δ is a parameter reflecting the weight that is put on the worst-off player's welfare, while $(1 - \delta)$ measures the weight that is put on the group's overall material payoff $\Sigma x_{j.}$

There is now a large body of experimental evidence in economics and psychology (Fehr and Schmidt, 1999; van Lange, 1999; Camerer, 2003) indicating that a substantial percentage of people are motivated by other-regarding preferences and that neither concerns for the well-being of others nor for fairness and reciprocity can be ignored in social interactions. In fact, social preferences in strategic interactions may play a decisive role for aggregate social and economic outcomes (Fehr and Gächter, 2000). However, the evidence also shows that there is considerable individual heterogeneity in social preferences: some people display little or no concern for their interaction partners, while others show strong social preferences. This heterogeneity in the strength of social preferences is a key reason why, in certain competitive environments, all individuals behave *as if* they were purely self-interested (Smith, 1962, 1982), while in strategic games the vast majority of individuals often deviate strongly from self-interested behavior. It is one of the great successes of social preference models (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Falk and Fischbacher, 2006) that they provide a parsimonious explanation of these puzzling facts.

The existence of social preferences does not mean that individuals make other-regarding choices no matter what costs they must bear. Rather, social preferences should be considered one important component in individuals' utility functions, implying that individuals with social preferences trade off other-regarding behavior with selfish goals: the more costly other-regarding behaviors are, the less likely it is that individuals will display such behaviors (Andreoni and Miller, 2002; Anderson and Putterman, 2006; Carpenter, 2007).

The fact that individuals are typically willing to trade off other-regarding actions with actions that maximize their material payoff is important because it enables us to model other-regarding behavior in terms of preferences or utility functions. This modeling further enables us to derive the implications and the limits of the impact of other-regarding preferences in interactive situations by means of game theoretic modeling.

MEASURING SOCIAL PREFERENCES

The main tools for eliciting social preferences are simple *one-shot* games such as the dictator game, the ultimatum game, or the third-party punishment game (see Box 15.2) that involve real monetary stakes and are played between anonymous interaction partners. A game is played one-shot if repeated play among the same two players is ruled out – that is, if the two players play the game with each other only once. In essence, an individual displays social preferences if he is willing to forgo his own material payoff for the sake of increasing or decreasing another individual's material payoff. For example, if an impartial observer (a "third party") in the third-party punishment game is willing to punish a greedy dictator who gives nothing to the recipient (see Box 15.2), and if the punishment is costly for the third party, his actions imply that he has a social preference.

Anonymity is important because it provides the conditions under which a baseline level of social preferences is observable. It is likely that face-to-face interactions change the strength and the pattern of social preferences, but this change can only be documented relative to the baseline. Moreover, a skeptic might argue that because face-to-face interactions inevitably involve an individual's reputation, the observed behaviors represent a combination of social preferences and *instrumental reputation-seeking*. The desire to

BOX 15.2

MEASURING SOCIAL PREFERENCES WITH ANONYMOUSLY PLAYED ONE-SHOT GAMES

Experimental games enable measurement of how much players are willing to sacrifice of their own economic payoff to increase or decrease the payoffs of others (Camerer, 2003; Fehr and Fischbacher, 2003). They provide a solid collection of empirical regularities from which the study of neural activity can proceed.

In a "dictator" game (Mikula, 1972; Kahneman *et al.*, 1986), one player – the dictator – is given a sum of money which he can allocate between himself and another player, the recipient. The dictator game measures a positive concern for the recipient's material payoff that is independent of the recipient's behavior, because the recipient can take no actions. Dictator allocations are a mixture of 50% offers and 0% offers (i.e., the dictator keeps everything), and a few offers in between 50 and 0%, but the allocations are sensitive to details of how the game is described (Camerer, 2003), the dictator's knowledge of who the recipient is (Eckel and Grossman, 1996), and whether the recipient knows that he is part of a dictator game (Dana *et al.*, 2006).

In an ultimatum game, the recipient can reject the proposed allocation (Güth *et al.*, 1982). If he rejects it, both players receive nothing. Rejections are evidence of negative reciprocity (Rabin, 1993), the motive to punish players who have behaved unfairly, or inequity aversion (Fehr and Schmidt, 1999), which is a distaste for unfair outcomes. The amount a recipient loses by rejecting a proposed allocation serves as a measurement of the strength of these motives. Offers of less than 20% are rejected about half the time; proposers seem to anticipate these rejections, and consequently offer on average approximately 40%. Cross-cultural studies, however, show that across small-scale societies the ultimatum offers are more generous when cooperative activity and market trade are more common (Henrich *et al.*, 2001).

In a third-party punishment game, two players, the dictator A and the recipient B, participate in a dictator game (Fehr and Fischbacher, 2004). A third player, the potential punisher C, observes how much A gives to B; C can then spend a proportion of his endowment on punishing A. This game measures to what extent "impartial" and "unaffected" third parties are willing to stick up for other players at their own expense, enforcing a sharing norm by punishing greedy dictators. Between 50% and 60% of the third parties punish selfish deviations from the equal split, suggesting that giving less than 50% in the dictator game violates a fairness norm. In principle, the third-party punishment option can be used to measure economic willingness to punish violation of *any* social norm (e.g., a violation of etiquette, breaking a taboo, or making a linguistic slur). In Fehr and Fischbacher (2004), for example, the third-party punishment game was used to document the existence of a "conditional cooperation" norm which prescribes cooperation conditional on others' cooperation.

In a trust or gift-exchange game, two players, A and B, each have an initial endowment. A first decides whether to keep his endowment or to send it to B. Then B observes A's action and decides whether to keep the amount he received or send some of it back to A. In a trust game (Camerer and Weigelt, 1988; Berg *et al.*, 1995), the experimenter doubles or triples A's transfer, whereas the back-transfer of player B is doubled or tripled in the gift-exchange game (Fehr *et al.*, 1993). Due the multiplication of A's transfer or of B's back-transfer, both players are better off collectively if A transfers money and B

BOX 15.2 *(Cont'd)*

sends back a sufficient amount. This situation mimics a sequential economic exchange in the absence of contract enforcement institutions. B has a strong incentive to keep all the money and send none to A; if A anticipates this behavior, however, there is little reason to transfer, so a chance for mutual gain is lost. Empirically, As invest about half of their endowment in the trust game and Bs repay about as much as player A invested (Camerer, 2003). Player As invest less than they do in risky choices with chance outcomes, however, indicating a pure aversion to social betrayal and inequality (Bohnet and Zeckhauser, 2004).

In a linear public goods game (Ledyard, 1995), players have a token endowment they can simultaneously invest in any proportion to a public project or keep for themselves. Investment into the public project maximizes the aggregate earnings of the group, but each individual can gain more by keeping the whole endowment. Typically, players begin by investing half their tokens on average (many invest either all or none). When the game is repeated over time, with feedback at the end of each decision period, investments decline until only a small fraction (about 10%) of the players invest anything. The prisoners' dilemma (PD) game is a special case of a public goods game, with two players and only two actions (cooperate or defect) for each player. When players are also allowed to punish other players at a cost to themselves, many players who invested punish the players who did not invest, which encourages investment and leads players close to the efficient solution in which everyone invests the whole endowment (Fehr and Gächter, 2002).

acquire a reputation that is profitable in future interactions is a purely self-regarding motive that has nothing to do with social preferences, i.e., it represents a confound. Therefore, the one-shot character and the anonymity in simple social preference experiments are crucial for the clean documentation of social preferences. Repeated interactions and a lack of anonymity are confounds that need to be eliminated if one is seeking a clean measure of social preferences.

A clean demonstration of social preferences also requires that an individual's action be independent of his belief about the opponent's action, because such beliefs affect behavior and therefore represent a confound. For this reason, the simultaneously played prisoners' dilemma (PD) game, which has often been used in the past to provide a measure of social preferences, is not appropriate for this purpose. The simultaneously played PD is a special case of a public goods game (see Box 15.2), and it is well known that many people are willing to cooperate in this game if they believe that their opponent will cooperate as well (Fischbacher *et al.*, 2001); however, if they believe that their opponent will defect, they will do so as well. Thus, defection in a simultaneous PD game does not necessarily indicate the absence of social preferences; it may merely be the result of pessimistic expectations about the other player's behavior.

Several theories of social preferences have been developed in the past 10–15 years (Andreoni, 1990; Rabin, 1993; Levine, 1998; Fehr and Schmidt, 1999; van Lange, 1999; Charness and Rabin, 2002; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006). All of these theories assume not only that subjects' utility functions contain their own material payoff as an argument, but also that non-pecuniary payoff elements, such as a concern for fairness, reciprocity, equality, or efficiency, enter into subjects' utility functions (see Box 15.1). In theories of *reciprocal fairness* (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006), for example, players are assumed to positively value other players' kind intentions, while negatively valuing their hostile intentions. Thus, if player A reduces B's payoff to his own benefit, a reciprocal player B will punish A, whereas if bad luck leads to a redistribution of income from B to A, a reciprocal player B will not punish (Blount, 1995). If, in contrast, a player is motivated by *inequity aversion* (Fehr and Schmidt, 1999), i.e. a dislike of unequal outcomes *per se*, bad luck will induce player B to take action to redistribute income (Dawes *et al.*, 2007). Likewise, some theories postulate an individual's desire to increase the economic welfare of the group they belong to (van Lange, 1999; Charness and Rabin, 2002), to experience a *warm glow* from altruistic giving to worthy causes (Andreoni, 1990), or to maintain a positive *social image* (Benabou and Tirole, 2006).

Social preferences have also been observed in experiments with relatively high stake levels (Hoffman *et al.*, 1996; Slonim and Roth, 1998; Cameron, 1999; Fehr *et al.*, 2002). Surprisingly, an increase in the amount at stake had no or only small effects on subjects' behavior. For example, Cameron (1999) conducted ultimatum games in Indonesia where subjects in the high-stake condition could earn the equivalent of three months' income in the experiment. She observed no effect of the stake level on proposers' behavior, and a slight reduction in the rejection probability when stakes were high.

Research has also documented only relatively small cross-cultural differences in social preferences in student populations from diverse Western countries (Roth *et al.*, 1991). However, large cross cultural differences have been observed across different small-scale societies, indicating that large variations in the cultural and institutional features of societies might lead to very different social preferences (Henrich *et al.*, 2001, 2006).

There is surprisingly little evidence, however, on the intra-personal stability of social preferences. The most convincing evidence comes from van Lange and co-authors (van Lange *et al.*, 1997; van Lange, 1999), who measured the social value orientation of a large number of subjects in a series of Dictator games – a technique which became known as the ring test (McClintock and Liebrand, 1988), where the data enable a graphic representation of social preferences on a circle ("ring"). He found a relatively large intra-personal stability, but more studies on intra-personal stability would certainly be desirable. Such replication is important in view of the current tendency to bring genetics to social preferences research (Wallace *et al.*, 2007). Research on the genetics of social preferences will require persuasive demonstrations of intra-personal stability.

ANTICIPATING OTHERS' SOCIAL PREFERENCES

Individuals with social preferences behave differently compared with those who only care about their material payoffs. Many individuals, however, are also aware of other people's social preferences. Even completely egoistic subjects often know that other people have social preferences, and this knowledge may cause them to change their behavior significantly. This fact is nicely illustrated in a recent paper by Spitzer *et al.* (2007), in which the same individuals played the proposer in a dictator game and the proposer in a punishment game similar to the ultimatum game. In both games, the proposer was given 100 money units and could transfer as much as he wanted to the recipient (responder). It is well known that the social norm in such games requires transferring half the money to the recipient (Fehr and Fischbacher, 2004). The crucial difference between the two games is that the responder can punish the proposer for unfair transfers in the Punishment game, while no punishment is possible in the dictator game. In the punishment game, the responders strongly punished transfers below the equal split; this led to a strong increase in average transfers, but there was considerable heterogeneity in response to the punishment threat (see Figure 15.1). Many proposers anticipated responders' punishment behavior and made much higher transfers in the punishment game right from the beginning, while some proposers first had to experience punishment before they increased their transfers relative to the dictator game. Moreover, individual differences in the transfer increase across conditions correlate with 0.5 with individuals' Machiavelli score, a measure of selfishness and opportunism. The score is based on a questionnaire (Christie and Geis, 1970) in which the subjects indicate their degree of agreement with statements such as "It's hard to get ahead without cutting corners here and there" and "The best way to deal with people is to tell them what they want to hear".

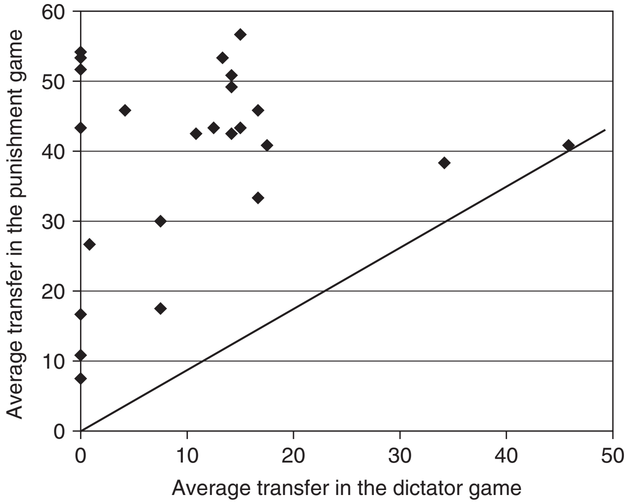


FIGURE 15.1 **Behavioral changes induced by the punishment threat**. The figure documents each proposer's average behavior in a dictator and in a punishment game. Each data point represents one individual. All 24 subjects transfer on average a higher amount to the recipient in the punishment game, indicating that all of them seem to be aware of the punishment threat. Eight of the 24 subjects even give zero or close to zero in the dictator game, indicating that they care only for their own payoff, but they transfer substantial amounts in the punishment game, suggesting that they anticipate that the recipients' social preferences give rise to a credible punishment threat. In fact, transfers below 50% were strongly punished, rendering such transfers unprofitable. A quarter of the subjects give on average even slightly more than 50% to ensure escaping punishment.

The anticipation of punishment driven by social preferences has also been shown to strongly increase cooperation in public good games (Fehr and Gächter, 2002) and contract enforcement games (Fehr *et al.*, 1997). If the members of a group or the contracting partners are given a costly punishment opportunity, compliance with cooperation norms and contractual obligations is much higher. Typically, the increase in compliance occurs immediately after the subjects are given punishment opportunities, indicating that some subjects anticipate the punishment threat instantaneously. However, some subjects have to learn the hard way – they only increase their compliance after having actually experienced punishment by others.

In situations involving trust, anticipation of the partner's social preferences is equally important because subjects with preferences for, say, reciprocity are more likely to repay trust. However, social preferences also influence the trust decision itself. This has been nicely demonstrated by Bohnet and Zeckhauser (Bohnet and Zeckhauser, 2004; Bohnet *et al.*, 2008), who conducted binary trust games in which they elicited subjects minimum acceptance probability (MAP). In their trust game, the trustors could either choose a sure option that gave both players a payoff of 10, or they could trust, which involved the risk that the trustee would not honor their trust. The trustee could reciprocate trust, giving both players a payoff of 15, or the trustee could defect, resulting in a payoff of 8 for the trustor and 22 for the trustee.

Bohnet and Zeckhauser elicited a trustor's MAP, which is the minimum probability for which the trustor is willing to make the trusting move, in two different conditions: (1) the trust game; (2) a risky dictator game identical to the trust game except that a computer mechanism forces a decision upon the trustee. Thus, in the risky dictator game the computer decides, according to an unknown predetermined probability p*, whether the trustee's trust is honored, but a human "trustee" collects the resulting earnings. In the trust game, the distribution of the trustees' actions determines p*; it is also unknown to the trustors. Announcing a MAP that is below p* is tantamount to choosing "trust" in the trust game with a randomly assigned trustee, while if the MAP is above p* the subject prefers the sure payoff S. The same procedure applies to the risky dictator game, except that the experimenters predetermine p*.

Interestingly, the trustors' MAPs are substantially higher in the trust game compared to the risky dictator game, indicating that the source of the involved risk affects the decision to trust (Bohnet *et al.*, 2008): if the trustee's choices determine the risk, the trustor is less willing to "trust", indicating more than just aversion against risk but also aversion against being betrayed by another human being. Bohnet *et al*. call this phenomenon "betrayal aversion." While they document that betrayal aversion is a robust feature across several different cultures, including the US, the United Arab Emirates, and Turkey, they also find substantial cross-cultural differences, with Brazil and China exhibiting much less betrayal aversion than the US or Turkey.

EXPLORING THE NEURAL CIRCUITRY OF SOCIAL PREFERENCES – METHODOLOGICAL CONCERNS

The rapid development of non-invasive brain-imaging and brain-stimulation methods now makes it possible to examine the neural networks involved in behavioral expressions of social preferences in humans. The combination of neuroscientific methods with interactive games in an attempt to study the neural processes behind social preferences categorically requires the use of games which actually allow the researcher to measure these social preferences. Games in which an individual interacts repeatedly with the same partner – a repeated PD, for example – are clearly inappropriate tools, because the behaviors in such games incorporate much more than just social preferences. Strategic sophistication plays a role in these games, and it is never clear whether a player responds to his opponent's past behavior or if he wishes to affect the interaction partner's future behavior. It is impossible to infer the players' motives cleanly using these games.

The best method for studying social preferences is to confront the experimental subject with a series of one-shot games in which the subject faces a different partner in every trial. In addition, the game should NOT be a simultaneous move game (e.g. the simultaneous PD), but should be played sequentially with the target subject being the second-mover who is informed about the first mover's choice. The second-mover knows the first mover's choice in the sequentially played PD, for example, and thus has no need to form expectations about the first mover's behavior. If the second mover makes a cooperative choice in response to the first-mover's cooperative action, this is a clean expression of a social preference because the second-mover gives up material payoff in order to cooperate, and beliefs about the first mover's possible actions do not confound the choice.

However, the implementation of a series of one-shot interactions poses a serious problem, because each subject in the brain scanner or under transcranial magnetic stimulation (TMS) needs to face a large number of other subjects. The temptation to deceive the subjects and to confront them with fabricated choices is therefore quite strong in this case – a strategy that may backfire in the medium or long run, because it undermines the experimenter's reputation. It is not always sufficiently acknowledged that one of the most important assets of a laboratory is its credibility and its reputation for being honest with the subjects. If subjects come to an experiment with the suspicion that the experimenter says "A" but in fact does "B," the experimenter loses control. To illustrate this point, suppose that subjects in a dictator game don't believe that the recipient in fact exists – that is, they believe that any money given to the recipient goes in fact to the experimenter. It is highly likely that suspicious subjects will behave more selfishly, and therefore the behavioral data will overstate the extent of selfishness. A possible way out of this dilemma is to confront the subjects in the scanner with choices that the interaction partners made in previously played identical games. This strategy has been implemented in de Quervain *et al.* (2004). In this study, the subjects in the behavioral pilot for the scanning study were asked at the end whether their choices could again be "used" for another study, and if they were used then the subjects indeed received the payments associated with their choices a second time. Thus, this strategy avoids deceiving the subjects in the scanner about the existence of their interactions partner and still allows the conduct of many one-shot games in the brain scanner.

Another solution to this problem is possible in the case of direct current stimulation (tDCS). tDCS induces changes in cortical excitability by means of a weak electrical field applied transcranially, which de- or hyperpolarizes neuronal membranes to a sub-threshold level. Anodal tDCS increases, while cathodal tDCS decreases, excitability (Nitsche and Paulus, 2001). It has been demonstrated that the neurophysiological and functional effects of tDCS are fairly restricted to the area under the electrodes (Nitsche *et al.*, 2003, 2007). A key feature of tDCS is that it is inexpensive and can be simultaneously applied to many subjects who interact in a laboratory environment (Knoch *et al.*, 2007). Thus, in principle, tDCS can be applied to a group of, say, 20 subjects simultaneously, with each of them playing one one-shot game with the other 19 subjects. Therefore, tDCS could prove to be a non-invasive brain stimulation method that revolutionizes neuroeconomics because it greatly enhances data collection efficiency and enables brain stimulations in whole groups of interacting subjects.

Another problem concerns the inferences that can be drawn from neuroimaging data in social preference tasks. In principle, subjects' choices in simple interactive games reveal social preferences if they deviate from the choices that maximize a subject's monetary payoff in particular ways – for example, by sending back money to the trustor in a one-shot anonymous trust game. The neural network activated during such choices thus reveals the neural circuitry of social preferences. We are frequently tempted, however, to reverse the inference process by inferring motivation and cognitive mechanisms from neuroimaging data. Our trust in such reverse inferences is justified if there is prior knowledge about the selectivity of the brain activation (Poldrack, 2006): if existing research has documented that the activated brain area used to infer the cognitive process is typically active when these cognitive processes occur, we can have more trust in such reverse inferences. Furthermore, trust in reverse inferences is higher if additional data, such as data, on mood, satisfaction, or response-time data, are available to bolster the reverse inference. For example, if activation in the ventral striatum is taken as evidence for expected rewards, it is important to have additional data available that support the hypothesis that subjects had a rewarding experience. Likewise, if activation in the amygdala is taken as evidence for fear, it is necessary to have other measures (such as skin conductance measures, self-report measures of fear, etc.) that support the fear hypothesis.

THE NEUROBIOLOGY OF OTHER-REGARDING PUNISHMENT BEHAVIOR

The Role of Prefrontal Cortex and the Anterior Insula

The readiness to reduce other people's income is a key feature of social preferences. This readiness may be triggered by the desire to punish unfair intention, to punish unfair people, or to re-establish equality in payoffs. The first neuroeconomic study of punishment behavior examined the responder in a series of one-shot ultimatum games with fMRI during the decision phase of the experiment (Sanfey *et al.*, 2003; see also Chapter 6 of this volume). This study reports activation of bilateral dorsolateral prefrontal cortex (DLPFC), bilateral anterior insula (AI), and the ACC in the contrast between "unfair – fair" offers. In addition, the higher the activation of right AI, the more likely a subject is to reject an unfair offer, suggesting that AI activation may be related to the degree of emotional resentment of unfair offers. Due to the proposed role of ACC in conflict monitoring (Botvinick *et al.*, 2001), ACC activation in this task may reflect the motivational conflict between fairness and self-interest when facing unfair offers. Finally, DLPFC activation may represent the cognitive control of the emotional impulse to reject unfair offers. A second fMRI ultimatum game study (Golnaz *et al.*, 2007) found increased AI activation (relative to a resting baseline) during trials with rejected unfair offers. However, this study did not find AI activation in the contrast between unfair – fair offers; instead, in a comparison of trials with unfair rejected offers with unfair accepted offers, these authors found increased left AI activation. In addition, the right VLPFC was more activated (relative to a resting baseline) when unfair offers were accepted, which may indicate that this region down-regulates the resentment associated with unfair offers. Consistent with the hypothesis that right RVPFC down-regulates AI, the study also found a negative correlation between activity in right VLPFC and left AI during trials in which unfair offers were accepted.

The two studies mentioned above are consistent with the idea that right DLPFC and VLPFC are involved in the cognitive control of the impulse to reject unfair offers. If this view is correct, an exogenous down-regulation of DLPFC activity reduces the control of this impulse and should therefore increase the rejection rate. Knoch *et al.* (2006a) examined this hypothesis by reducing the activation in right and left DLPFC with low-frequency TMS. Contrasting with the hypothesis, however, this study found that TMS of right DLPFC increases the *acceptance* rate of unfair offers relative to a placebo stimulation (from 9% to 44%) and relative to an active stimulation of left DLPFC, which left acceptance rates unaffected (see Figure 15.2). Another study (van 't Wout *et al.*, 2005) also report's a tendency for right DLPFC stimulation to increase the acceptance rate of unfair offers relative to a sham stimulation. Low-frequency TMS of right DLPFC did not have a significant effect in this study, but this may be due to the small number of subjects (Seven) and the fact that the authors implemented a within-subject design. It is well known that TMS, in particular TMS of the PFC, can be irritating for subjects, while sham stimulation is not. There is therefore a high probability that in a within-subject design the participants know whether they are receiving sham stimulation or real stimulation. Further evidence for a causal role of right DLPFC in responder behavior comes from a study that uses direct current stimulation (tDCS) to reduce DLPFC activity (Knoch *et al.*, 2007). This study found a substantial and significant positive effect of tDCS on the acceptance rate for unfair offers.

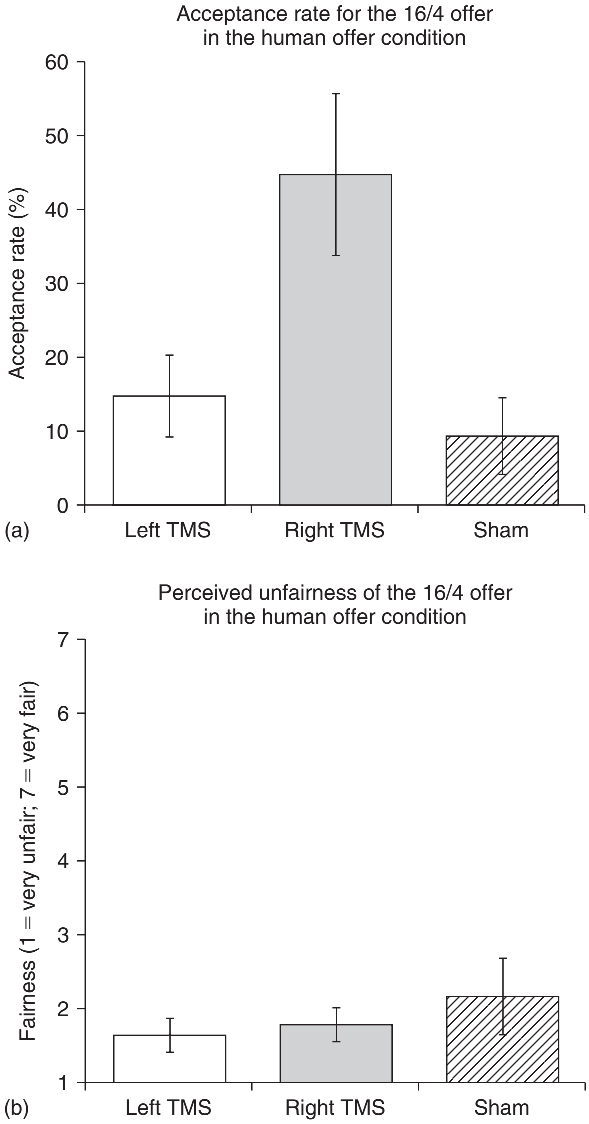


FIGURE 15.2 Acceptance rates and fairness judgments (means ± s.e.m.) related to the most unfair offer of CHF 4 in the human offer condition. (a) Acceptance rates across treatment groups. Subjects whose right DLPFC is disrupted exhibit a much higher acceptance rate than those in the other two treatment groups. (b) Perceived unfairness across treatments (1 = very unfair; 7 = very fair). Subjects in all three treatment groups perceive an offer of 4 as very unfair, and there are no significant differences across groups.

These findings suggest that right DLPFC activity is crucial for the behavioral implementation of fairness motives (i.e., rejection of unfair offers), and not for the implementation of selfish choices, as the previous studies hypothesized. The facts in Knoch *et al.* (2006a) further support this interpretation. If a computer-generates low offers, the effect of TMS to the right DLPFC is strongly mitigated and insignificant. It is known that computer-generated low offers are viewed as much less unfair (Blount, 1995) than if a human proposer makes low offers, i.e., the fairness motive for rejecting a low offer is weaker for a computer-generated offer. Therefore, if the implementation of a fairness motive requires activation in the right DLPFC, a weaker fairness motive is likely to be associated with a lower recruitment of DLPFC, and thus a lower chance of disrupting the implementation of the fairness motive with TMS.

Two different phenomena could cause the lower ability to implement the fair choice. First, subjects' fairness judgments may have changed due to the disruption of right DLPFC, meaning that they may no longer view low offers as unfair and therefore will be less likely to reject them. Second, subjects may still view low offers as unfair, but may no longer be able to resist the selfish temptation to accept unfair offers. The evidence is consistent with the second hypothesis, because TMS to the right DLPFC (see Figue 15.2) failed to affect subjects' fairness judgments. Thus, they continue to view low offers as very unfair but nevertheless accept them at a much higher rate. Response-time data further support the hypothesis that TMS to the right DLPFC reduces the ability to resist selfish temptations. If subjects face a fair offer, there is no conflict between self-interest and fairness, and it takes them on average slightly more than 3 seconds to accept the offer. If subjects face an unfair offer and receive sham stimulation or TMS to the left DLPFC, it takes them roughly 6 seconds to accept an unfair offer, indicating that the conflict between self-interest (acceptance) and fairness (rejection) delays the decision. If, however, subjects face an unfair offer and receive TMS to the right DLPFC, they accept unfair offers as quickly as they do fair offers. Thus, in terms of response time, subjects behave as if there is no longer a conflict between self-interest and fairness, consistent with the hypothesis that they are no longer or less able to resist the selfish temptation.

What role does the AI play in this interpretation? The AI was linked to the emotional representation of the cost of purchasing a consumer product in a recent paper (Knutson *et al.*, 2007). In this study, higher insula activation in response to cost information correlated with a reduction in the probability that subjects would subsequently buy the consumer product. The AI activation has been linked to the emotional resistance to accept an unfair offer in the ultimatum game, or, in other words, to the neural representation of the emotional cost of accepting an unfair offer. Perhaps the disruption of right DLPFC reduces a subject's ability to process this cost information or to integrate this cost information with the monetary benefits that are associated with the acceptance of unfair offers.

Previous research suggests that the ventromedial prefrontal cortex (Brodman areas BA 10, 11) is involved in the integration of separate benefits and costs in the pursuit of behavioral goals (Ramnani and Owen, 2004). Support for this view also comes from the study of de Quervain *et al.* (2004), which reports BA 10 and 11 activation in the contrast between a costly punishment condition and a costless punishment of trustees who defected in the trust game. Further corroboration is found both in an fMRI study of charitable donations (Moll *et al.*, 2006), which documents VMPFC activation (BA 10, 11, 32) in the contrast between altruistic decisions involving costs and no costs, and in a purchasing task where VMPFC activity seems to integrate the value of consumer products and their costs (Knutson *et al.*, 2007). In fact, VMPFC in the purchase task is positively correlated with subjects' assessment of the *net value* of the product (i.e. value − price), which provides nice support for the "integration hypothesis." More generally, recent evidence indicates that the VMPFC is involved in emotional processing and moral judgment (Moll *et al.*, 2005; Koenigs *et al.*, 2007); lesions to VMPFC are associated with poor choices in various situations (Damasio, 1995; Bechara *et al.*, 1997) which require integrating costs and benefits. These studies suggest a general role of VMPFC in integrating emotional feelings about costs and benefits, regardless of whether these choices involve economic consumption goods or "non-economic" goods such as the subjective value of rejecting an unfair offer.

In view of the important role of VMPFC in emotional processing and in integrating emotional feelings about costs and benefits, it seems possible that low-frequency TMS of right DLPFC induces an impairment in the integration of the emotional cost of accepting an unfair offer. Such an impairment could be caused by possible network effects of TMS that diminish the functioning of the VMPFC. Network effects of TMS have been shown in several studies (Wagner *et al.*, 2007); a recent PET study (Eisenegger *et al.*, 2008) shows that low-frequency rTMS of right DLPFC increases blood flow in the right DLPFC *and* the right VLPFC if subjects perform no task during PET. Of course, these network effects should ideally be studied during the task under consideration – in our case, the responders' decision in the ultimatum game – because the TMS effects are likely to be different depending on whether a brain area is recruited during a task or not.

Social Preferences and Reward Circuitry

Social preference theories assume that material payoffs are transformed into subjective payoffs that give rise to the altruistic, fairness, and reciprocity related behaviors described in Box 15.2. While this idea also has a long tradition in psychology (Thibaut and Kelley, 1959), psychologists rarely developed precise formal theories, such as those presented in Box 15.1, that could be plugged into game theoretic models. Although social preference theories do not make assumptions about the hedonic processes associated with these behaviors (because they rely on inferred decision utilities), a plausible interpretation of these theories is that subjects in fact derive higher hedonic value from the mutual cooperation outcome. Indeed, questionnaire evidence (M. Kosfeld, E. Fehr, and J. Weibull, unpublished) supports the view that mutual cooperation in social exchanges has special subjective value, beyond that associated with monetary earnings (Fehr and Camerer, 2007). An obvious question is therefore whether we can find neural traces of the special reward value of the mutual cooperation outcome. A neuroimaging study (Rilling *et al.*, 2002) reports activation in the ventral striatum when subjects experience mutual cooperation with a human partner compared to mutual cooperation with a computer partner. Despite the fact that the monetary gain is identical in both situations, mutual cooperation with a human partner is associated with higher striatal activity, consistent with the reward hypothesis, given that substantial evidence from other studies with primary and secondary rewards indicates that the anticipated rewards activate the striatum.

Social preference theories also predict that subjects prefer punishing unfair behavior, such as defection in public good and PD games, because leaving an unfair act unpunished is associated with higher disutility than bearing the cost of punishing an unfair act. In this view, it is natural to hypothesize that the act of punishing defection involves higher activation of reward circuitry. A study using PET (de Quervain *et al.*, 2004) examined this hypothesis in the context of a trust game in which the trustor had a punishment opportunity after he had observed the trustee's choice. This study showed that the dorsal striatum (caudate nucleus) is strongly activated in the contrast between a real punishment condition (in which the assignment of punishment points hurt the defector in economic terms) and a symbolic punishment condition (in which the assignment of punishment points did not reduce the defector's economic payoff). In another study (Singer *et al.*, 2006), subjects first played a sequential PD with (confederate) fair and unfair opponents. The focal subjects were then scanned (using fMRI) when a slight pain – an electrical shock – was administered either to themselves or to confederate partners who behaved fairly or unfairly. Both men and women exhibited empathic responses in anterior cingulate and AI when the fair partner received pain. However, only men report a higher desire for revenge against unfair partners, while also exhibiting activation in the nucleus accumbens (NAcc) and orbitofrontal cortex (OFC) when unfair partners are shocked. Male revenge-desire ratings across subjects are also correlated with the estimate of NAcc activity, consistent with the view that there is reward value for men in observing the punishment of unfair partners.

If the punishment of unfair behavior is rewarding, why did the authors of the first fMRI ultimatum game study (Sanfey *et al.*, 2003) fail to observe reward-related activity in the striatum? The answer to this question is informative, because it highlights the role of experimental designs and the contrasts that they permit to be computed. In brief, this study was not designed to measure reward-related activity. In order to do so, similar treatment conditions to those in the DeQuervain study would have been necessary – a condition where a rejection reduces the proposer's payoff to zero, and one where a rejection only slightly reduces the proposer's payoff or does not do so at all. The contrast between these two conditions provides information about the encoding of rewards through punishment in the striatum.

Further evidence that decisions involving social preferences are associated with activity in reward circuitry comes from fMRI studies of charitable donations (Moll *et al.*, 2006; Harbaugh *et al.*, 2007; see also Chapter 19 of this volume) and reaction to offers in an ultimatum bargaining game (Golnaz *et al.*, 2007). Both ventral tegmental (VTA) and striatal areas are activated by receiving money and non-costly donations, indicating that "giving has its own reward" (Moll *et al.*, 2006). Across subjects, those who made more costly donations also had more activity in the striatum. Decisions to donate, whether costly or not, activate the subgenual area, a region densely connected with mesolimbic, dopaminergic, and serotonergic pathways, and implicated in social attachment mechanisms and in regulating the release of the neuromodulator oxytocin via the anterior hypothalamus. In Harbaugh *et al.* (2007), subjects faced two conditions – a forced donation and a voluntary donation condition. In the former, subjects passively observed that money was transferred either to their account or to the charities account. In the voluntary condition, the subjects could decide whether to accept these monetary transfers. Subjects reported higher satisfaction in both the forced and the voluntary condition if they themselves received more money and if the charity received more money (controlling for the subject's cost of this transfer). Moreover, activations in dorsal and ventral striatum in both conditions positively correlated with the money that went to the charity and to the subjects themselves. A recent ultimatum game study (Golnaz *et al.*, 2007) provides evidence suggesting that the fairness of a bargaining offer – controlling for the absolute size of the monetary gain – is associated with activations in the ventral striatum. The same dollar bargaining offer of, say, $5 elicits higher striatal activation if the offer represents a fair share (say, 50%) of the amount being bargained over, compared to when that dollar offer represents a small share (say, only 15%). Finally, a recent paper (Fliessbach *et al.*, 2007) provides evidence that social comparison of monetary rewards is associated with reward-related activity in the human ventral striatum. Controlling for a subject's own absolute income level, this study shows that variations in the comparison subject's income lead to variations in ventral striatum activity. This finding is interesting, as several social preference theories (e.g. the theory of Fehr and Schmidt, 1999) are based on the assumption of social comparison processes.

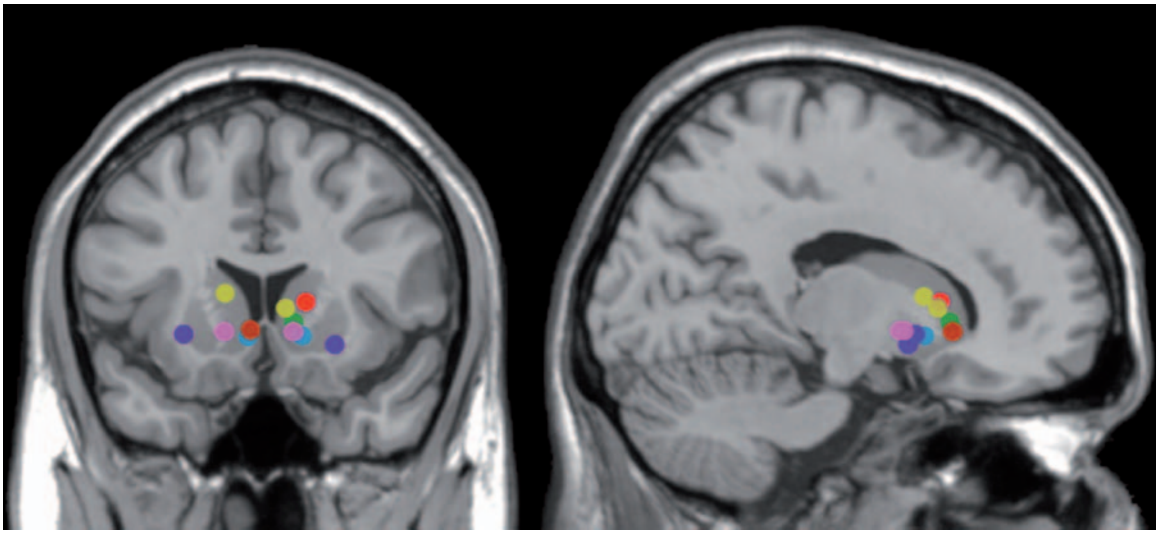
In view of the above results, a comparison of activations in the striatum when subjects themselves expect or receive rewards to when they obtain social rewards through altruistic, fairness, or reciprocity related behaviors is informative. Figure 15.3a shows typical activations in case of own material rewards, while Figure 15.3b illustrates the activations observed in studies of social rewards. These figures clearly show a substantial overlap between these areas of activation, consistent with the hypothesis that social preferences are similar to preferences for one's own rewards in terms of neural activation.

The above evidence is consistent with the view that both costly pro-social acts of charitable donation and punishment of unfair behaviors are rewarding. However, the hedonic interpretation of social preference theories also implies that such acts occur *because* they are rewarding. Evidence for causality is also important for moving from correlation to causality, and because some studies suggest that a unitary system, centered on the striatum, might fully process both monetary gains *and* losses (Seymour *et al.*, 2007; Tom *et al.*, 2007). If the striatum is generally more active for positively *and* negatively valued goods, the mere fact that studies show higher fMRI BOLD responses for costly altruistic acts may indicate a costly experience rather than a rewarding one. But if it could be shown that higher activations in the striatum *imply* a higher willingness to act altruistically, the case for the reward interpretation would be strengthened considerably – since it is implausible to observe this relation between striatum activation and altruistic acts if striatum activation represents the cost of the act rather than its reward value.

While neuroimaging data do not allow causal inferences, it is possible to move towards causality by using neural activity in one treatment to predict choice behavior in another ("out of treatment" forecasting). In DeQuervain *et al.* (2004), for example, individual differences in caudate nucleus activation when punishment is costless for the punisher predict how much individuals will actually pay for punishment when it is not costless. Likewise, in Harbaugh *et al.* (2007), individual differences in striatal activity in the condition in which donations are forced predict subjects' willingness to donate money to charities in the condition where donations are voluntary. Moreover, individual differences in the extent to which social comparison affects activity in the ventral striatum predict subjects' willingness to reciprocate kind or unkind acts (Fliessbach *et al.*, 2007). These results further support the reward interpretation of social preferences, which in turns provides support for the hypothesis of a common neural currency of socially-preferred and other primary and secondary rewards (Montague and Berns, 2002).

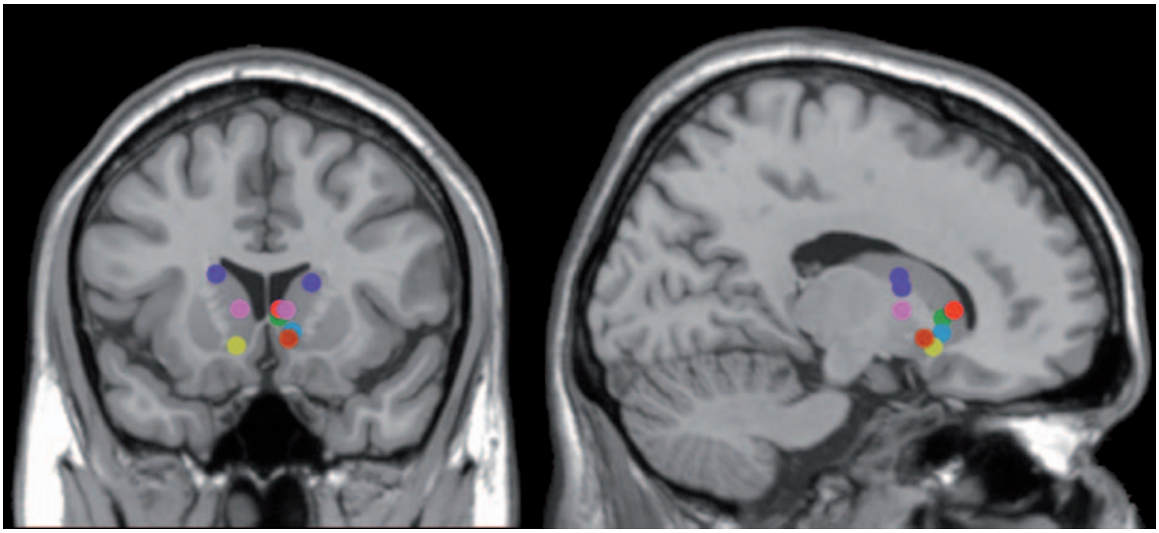
How Does the Brain Anticipate Social Punishment?

While several neuroeconomic studies already examine the brain circuitry of punishers, very little work has been conducted on the brain circuitry of those who face the possibility of punishment for unfair behaviors. A recent study (Spitzer *et al.*, 2007) compared subjects' behavioral and brain responses in a punishment game and in a dictator game. The behavioral results in this study were discussed above and illustrated in Figure 15.1. In terms of brain activation, the study was based on two hypotheses. First, previous work (Kringelbach, 2005) strongly suggests that lateral OFC is involved in processing of punishing stimuli that can induce behavioral changes. When self-interest tempts subjects to violate a fairness norm while facing the threat of punishment, they are in exactly this situation: they face a punishing stimulus if they violate the norm, and they subsequently have the chance to comply with the norm in response to punishment. Second, the TMS and tDCS results of Knoch *et al.* (2006a) Knoch *et al.*, (2007) suggest that the right DLPFC is involved in the control of self-interested impulses. Thus, if subjects' immediate self-interest tempts them to violate a fairness norm, it seems plausible to hypothesize that the right DLPFC is recruited in order to control this impulse in the service of a subjects'



(Bhatt and Camerer, 2005; Delgado *et al.*, 2003; Knutson *et al.*, 2001a; Knutson *et al.*, 2001b; O'Doherty J, 2004; Preuschoff *et al.*, 2006; Tricomi *et al.*, 2004)

(a)



(De Quervain *et al.*, 2004; Fliessbach *et al.*, 2007; Harbaugh *et al.*, 2007; Moll *et al.*, 2006; Rilling *et al.*, 2002; Rilling *et al.*, 2004; Singer *et al.*, 2006)

(b)

FIGURE 15.3 Parallelism of rewards for oneself and for others: Coronal and sagittal representation of brain activations in (a) a sample of seven studies of monetary reward processing and (b) seven studies of social reward. Coronal sections: All activations are projected onto Tailarach coordinate $y = 15$; voxels of peak activation in original studies range from $y = 4$ to $y = 24$. Sagittal sections: All activations are projected onto Talairach coordinate $x = 16$; voxels of peak activation in original studies range from $x = -26$ to $x = 26$. Pictures are constructed by Christoph Eisenegger.

enlightened self-interest, which includes the goal of avoiding punishment for norm violations.

The brain-imaging results in Spitzer *et al.* (2007) support both hypotheses – bilateral OFC and bilateral DLPFC are more strongly activated during subjects' choices in the punishment game compared to the dictator game. In addition, individuals with stronger activation in the right DLPFC and the left OFC, in the decision phase show stronger increases in norm compliance when they face the punishment threat, lending further support for a crucial involvement of these brain areas in norm compliance. The social nature of the punishment threat also modulates activations in both the right DLPFC and the left OFC, because the activation in these brain areas is stronger if the responder can punish compared to a situation where subjects face a punishing computer. Finally, the contrast between social punishment and computer punishment shows activation in the right insula and the left STS, suggesting that the brain makes different computations if subjects face punishment from the victim of a *social* norm violation compared to punishment "by nature" (in this case, a computer).

The above results might be useful in the study of people with antisocial personality disorder who have been shown to display deficient lateral OFC and insula activation in fear-conditioning paradigms (Veit *et al.*, 2002; Birbaumer *et al.*, 2005). The experimental paradigm in Spitzer *et al.* (2007) could be used simultaneously to study the behavioral responses of "psychopaths" together with their brain responses in situations involving the temptation to violate a social norm.

THE NEUROBIOLOGY OF TRUST AND TRUSTWORTHINESS

Social preferences models predict that trusting other individuals by making investments that may not be repaid is not just a decision involving monetary risk. Reciprocal and inequity-averse subjects derive a special dis-utility from betrayal of trust, along with the associated economic loss; this is consistent with behavioral studies (Bohnet and Zeckhauser, 2004) indicating a pure aversion to social betrayal. The first evidence that the brain distinguishes between social trust and monetary risk-taking came from Kosfeld *et al.* (2005), who infused the synthetic neuropeptide oxytocin (OT) intranasally to players in a trust game. OT-infused players were more trusting than a placebo control group, although their beliefs about the chances of being repaid were not higher and OT did not affect risk-taking in a pure risk condition. Thus, OT seems to limit the fear of betrayal in social interactions, consistent with animal evidence that it inhibits defensive behavior and facilitates maternal behavior and pair-bonding (Insel and Young, 2001). The hypothesis that oxytocin's dampening effect on the fear of betrayal might be due to a reduced activation of the amygdala is consistent with a study (Kirsch *et al.*, 2005) showing that OT dampens amygdala activity and its connections to the brainstem if subjects view emotionally arousing pictures. In fact, the study of Baumgartner et al. (2008) which combines the intranasal administration of OT in a trust game with fMRI, shows that OT affects trusting behavior only in those situations where OT also dampens amygdala activity. Amygdala involvement has also been shown to occur in assessing the trustworthiness of faces (Winston *et al.*, 2002; Adolphs *et al.*, 2005).

Since trust decisions are also likely to involve perspective-taking, they should also activate areas implicated in theory-of-mind tasks, such as the paracingulate cortex and the posterior superior temporal sulcus (pSTS) (Frith and Frith, 2003). One of the earliest neuroeconomic studies (McCabe *et al.*, 2001) reported activation of the paracingulate cortex in a trust game when subjects played against another person compared to a computerized opponent. Another study found that pSTS is activated simply by showing the faces of intentional cooperators compared to non-intentional agents (Singer *et al.*, 2004).

In repeated trust games, a player learns about his opponent's choices, meaning that the opponent acquires a reputation. In game theory, this reputation is defined as the subjective probability that the opponent is the type of player who prefers to reciprocate trust. In this approach, players' preferences and their subjective beliefs are distinct concepts, and a rational player's beliefs are not colored by his preferences or his emotions towards the opponent. It is interesting to examine whether the brain also makes this distinction – i.e., whether the neural networks involved in hedonic preferences and emotional processing are distinct from those involved in assessing the opponent's reputation, or whether there is substantial overlap in these neural networks.

Preliminary evidence suggests that the latter is likely to be true. In one study (Singer *et al.*, 2004), players faced a series of cooperative and non-cooperative opponents in a sequential PD game. The authors found that simply displaying the faces of cooperative partners (relative to neutral faces) in a subsequent gender-assessment task activated striatal and emotion-related areas such as the amygdala, the insula, and the putamen. This suggests that a trustworthy person's face automatically triggers emotions and reward expectations, as if simply seeing another person's face activates its representation as a future exchange value.

Three other studies have demonstrated the importance of the striatum in learning the opponent's trustworthiness. In King-Casas *et al.* (2005), the activity in caudate nucleus signals whether the other player reciprocates an earlier move. The study of Baumgartner *et al.* (2008) also suggests a crucial role of the caudate nucleus when subjects learn about the trustworthiness of a population of trustees. Subjects who were given oxytocin did not change their trusting behavior after they received information that many trustees had betrayed their trust in previous interactions, whereas subjects who received placebo reduced their trusting behavior after this information. Moreover, OT caused a specific activity reduction in caudate nucleus, suggesting that the lack of trust adaption in subjects with OT may have been caused or modulated by the diminished recruitment of reward learning circuitry.

While several neuroimaging studies have examined the trustor, no work to date has studied the brain circuitry involved in the decision to repay trust. This represents an important research opportunity, because repaying trust truly involves social preferences if the game is played on a one-shot basis. Kosfeld *et al*. (2005) showed that oxytocin has no effect on the trustor's decisions, suggesting that the neural circuitry of trusting differs from the circuitry involved in repaying trust. The study by King-Casas *et al*. (2005) examined the trustee's brain activations, but used a repeated trust game in which the same partners interacted for ten periods. Thus, the trustee's decisions in this game can always be interpreted as a response to the trustor's previous choices or to the trustee's expectations about the trustor's future choices. It is therefore not possible to infer something about the neural networks specifically associated with the trustor's other-regarding preferences. Likewise, the simultaneously played PDs examined in Rilling *et al*. (2002, 2007) do not enable us to isolate the brain networks specifically involved in the processing of other-regarding choices, because it is unclear whether the PD players respond to their beliefs about the other player's likely behavior in the current period or to the other's previous choices.

CONCLUSIONS

Behavioral experiments show that many people exhibit other-regarding behaviors. Economists have developed precise mathematical theories that model concerns for reciprocal fairness, inequity aversion, and the whole group's material welfare. The precision of these mathematical models provides a deeper understanding of the motivational forces at work, enables researchers to derive aggregate implications of these forces in interactive games, and facilitates the design of new experiments that help in discriminating between models. Social preference theories and the knowledge about the relative strength of the different motivational forces in different environments can also be useful in interpreting neural data.

One emerging theme of the studies reviewed above is that social reward activates circuitry that overlaps, to a surprising degree, with circuitry which anticipates and represents other types of rewards. These studies reinforce the idea that social preferences for donating money, rejecting unfair offers, repaying trust, and punishing those who violate norms are genuine expressions of preference. The social rewards are traded off with subjects' economic self-interest, and the dorsolateral and ventromedial prefrontal cortex are likely to be crucially involved in balancing competing rewards. Treatments like oxytocin infusion and TMS disruption can also alter these processes, actually changing behavior in ways that are consistent with hypotheses derived from fMRI.

However, the knowledge about the neural networks involved in social preferences is still quite limited. In addition, very little is known about the causal impact of emotions (and the associated neural circuitry) on social preferences. Emotions must be induced experimentally in order to study their causal impact. Thus, improving the methods of emotion induction, in particular with respect to their power to affect behavior and their specificity, is an important task. Film clips used for emotion induction, for example, are a rather imprecise and non-specific method of emotion induction because films that, say, induce anger often also induce disgust. Moreover, I have yet to see a convincing induction of complex emotions such as guilt or shame; inducing these emotions would seem to be particularly important in the realm of social preferences.

We also know relatively little about why face-to-face communication – and the implicit emotional communication which occurs in face-to-face interactions – shapes other-regarding behaviors. For example, the equal split is offered in about 90–100% of the cases in face-to-face one-shot ultimatum games, as opposed to less than 50% in anonymous games. In a similar vein, I expect very high trust levels and very high back-transfers in one-shot trust games played face-to-face. Why does face-to-face interaction have such a powerful effect on other-regarding behaviors or beliefs about other-regarding behaviors? Standard economists may have an easy answer to this question. Face-to-face interactions inevitably remove anonymity, thus increasing the importance of reputation incentives.

However, I consider it highly unlikely that instrumental reputation-seeking alone can explain the vast behavioral changes induced by face-to-face interactions. Instead, I believe that subtle forms of emotional communication and anticipated emotional responses take place that change our social motives or put constraints on our selfish motives. Also, instrumental reputation-seeking – that is, the higher future material payoff conferred by a good reputation – is only one reason why humans care about reputation. Perhaps equally important, maybe even more so, is the fact that humans care about their reputation *per se*. Losing one's reputation is likely to be *per se* averse.

Research on the emotional underpinnings of social preferences, and the neural networks involved in "social preference emotions," may also be useful for economics, because it helps us to understand when social preferences are likely to affect behavior and when not. Such research may also help us gain deeper insights into the reasons for interindividual differences in social preferences. In this context, studies examining whether social preferences are stable attributes of individuals or stable attributes of situations are important. To date, very few studies examine the intra-individual stability of social preferences – a topic that is of great importance for economics. If social preferences are stable personality traits, it makes sense for organizations to invest resources into the screening of people; however, if social preferences are mainly an attribute of the environment, it makes sense to shape them, designing the environment appropriately.

Finally, I would like to argue strongly in favor of using a mix of methods in future neuroeconomic experiments. The previous almost exclusive emphasis on neuroimaging – unconstrained by computational models – is likely to yield sharply decreasing returns. We need a combination of methods that enable insights into the causes of behavior; TMS, tDCS, pharmacological interventions, computational models and techniques such as Dynamic Causal Modeling of neuroimaging data are crucial for future advances.

References

Adolphs, R., Gosselin, F., Buchanan, T.W. *et al*. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature* 433, 68–72.

Anderson, C.M. and Putterman, L. (2006). Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism. *Games Econ. Behav.* 54, 1–24.

Andreoni, J. (1990). Impure altruism and donations to public goods: a theory of warm glow giving. *Economic J.* 100, 464–477.

Andreoni, J. and Miller, J. (2002). Giving according to Garp: an experimental test of the consistency of preferences for altruism. *Econometrica* 70, 737–753.

Arrow, K.J. (1981). Optimal and voluntary income redistribution. In: S. Rosenfield (ed.), *Economic Welfare and the Economics of Soviet Socialism: Essays in Honor of Abram Bergson*. Cambridge: Cambridge University Press, pp. 267–288.

Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U. and Fehr, E. (2008). Oxytocin shapes the neural circuitry of trust adaptation in humans. *Neuron* 58, 639–650.

Bechara, A., Damasio, H., Tranel, D., and Damasio, A.R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science* 275, 1293–1295.

Becker, G.S. (1974). A theory of social interactions. *J. Political Economy.* 82, 1063–1093.

Benabou, R. and Tirole, J. (2006). Incentives and prosocial behavior. *Am. Econ. Rev.* 96, 1652–1678.

Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, reciprocity and social history. *Games Econ. Behav.* 10, 122–142.

Bhatt, M. and Camerer, C.F. (2005). Self-referential thinking and equilibrium as states of mind in games: fMRI evidence. *Games Econ. Behav.* 52, 424–459.

Birbaumer, N., Veit, R., Lotze, M. *et al*. (2005). Deficient fear conditioning in psychopathy: a functional magnetic resonance imaging study. *Arch. Gen. Psych.* 62, 799–805.

Blount, S. (1995). When social outcomes aren't fair – the effect of causal attributions on preferences. *Org. Behav. Hum. Dec. Proc.* 63, 131–144.

Bohnet, I. and Zeckhauser, R. (2004). Trust, risk and betrayal. *J. Econ. Behav. Org.* 55, 467–484.

Bohnet, I., Greig, F., Hermann, B., and Zeckhauser, R. (2008). Betrayal aversion – evidence from Brazil, China, Switzerland, Turkey, the United Arab Emirates and the United States. *Am. Econ. Rev.*, in press.

Bolton, G.E. and Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *Am. Econ. Rev.* 90, 166–193.

Botvinick, M.M., Braver, T.S., Barch, D.M. *et al*. (2001). Conflict monitoring and cognitive control. *Psychological Rev.* 108, 624–652.

Camerer, C.F. (2003). *Behavioral Game Theory – Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Camerer, C.F. and Fehr, E. (2006). When does "economic man" dominate social behavior? *Science* 311, 47–52.

Camerer, C. and Weigelt, K. (1988). Experimental tests of a sequential equilibrium reputation model. *Econometrica.* 56, 1–36.

Cameron, L.A. (1999). Raising the stakes in the ultimatum game: experimental evidence from Indonesia. *Econ. Inquiry.* 37, 47–59.

Carpenter, J. (2007). The demand for punishment. *J. Econ. Behav. Org.* 62, 522–542.

Charness, G. and Rabin, M. (2002). Understanding social preferences with simple tests. *Q. J. Economics* 117, 817–869.

Christie, R. and Geis, F. (1970). *Studies in Machiavellism*. New York, NY: Academic Press.

Damasio, A.R. (1995). *Descartes' Error: Emotion, Reason and the Human Brain*. New York, NY: Hayter Collins.

Dana, J., Cain, D.M., and Dawes, R.M. (2006). What you don't know won't hurt me: costly (but quiet) exit in dictator games. *Org. Behav. Hum. Dec. Proc.* 100, 193–201.

Dawes, C.T., Fowler, J.H., Johnson, T. *et al*. (2007). Egalitarian motives in humans. *Nature* 446, 794–796.

Delgado, M.R., Frank, R.H., and Phelps, E.A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8, 1611–1618.

Delgado, M.R., Locke, H.M., Stenger, V.A., and Fiez, J.A. (2003). Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cogn. Affect. Behav. Neurosci.* 3, 27–38.

de Quervain, D., Fischbacher, U., Treyer, V. *et al*. (2004). The neural basis of altruistic punishment. *Science* 305, 1254–1258.

Dufwenberg, M. and Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games Econ. Behav.* 47, 268–298.

Eckel, C. and Grossman, P. (1996). Altruism in anonymous dictator games. *Games Econ. Behav.* 16, 181–191.

Eisenegger, C., Treyer, V., Fehr, E., and Knoch, D. (2008). Time-course of "off-line" prefrontal rTMS effects – a PET study. *NeuroImage*, (in press).

Falk, A. and Fischbacher, U. (2006). A theory of reciprocity. *Games Econ. Behav.* 54, 293–315.

Fehr, E. and Camerer, C.F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* 11, 419–427.

Fehr, E. and Fischbacher, U. (2003). The nature of human altruism. *Nature* 425, 785–791.

Fehr, E. and Fischbacher, U. (2004). Third-party punishment and social norms. *Evol. Hum. Behav.* 25, 63–87.

Fehr, E. Fischbacher, U., and Tougareva, E. (2002). Do high stakes and competition undermine fairness? Evidence from Russia. Working Paper No 120, Institute for Empirical Research in Economics, University of Zurich.

Fehr, E. and Gächter, S. (2000). Fairness and retaliation: the economics of reciprocity. *J. Econ. Persp.* 14, 159–181.

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature* 415, 137–140.

Fehr, E. and Schmidt, K.M. (1999). A theory of fairness, competition, and cooperation. *Q. J. of Economics* 114, 817–868.

Fehr, E. and Schmidt, K. (2003). Theories of fairness and reciprocity: evidence and economic applications. In: M. Dewatripont, L.P. Hansen, and S.J. Turnovky (eds), *Advances in Economics and Econometrics*. Cambridge: Cambridge University Press, pp. 208–257.

Fehr, E., Kirchsteiger, G., and Riedl, A. (1993). Does fairness prevent market clearing – an experimental investigation. *Q. J. Economics* 108, 437–459.

Fehr, E., Gächter, S., and Kirchsteiger, G. (1997). Reciprocity as a contract enforcement device: experimental evidence. *Econometrica* 65, 833–860.

Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Econ. Letts.* 71, 397–404.

Fliessbach, K., Weber, B., Trautner, P. *et al*. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *Science* 318, 1305–1308.

Frith, U. and Frith, C.D. (2003). Development and neurophysiology of mentalizing. *Phil. Trans. R. Soc. Lond. B.* 358, 459–473.

Golnaz, T., Satpute, A.B., and Lieberman, M.D. (2007). The sunny side of fairness_ preference for fairness activates reward circuitry. *Psychological Sci.* in press.

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Org.* 3, 367–388.

Harbaugh, W.T., Mayr, U., and Burghart, D.R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* 316, 1622–1625.

Henrich, J., Boyd, R., and Bowles, S. *et al.* (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *Am. Econ. Rev.* 91, 73–78.

Henrich, J., McElreath, R., Barr, A. *et al*. (2006). Costly punishment across human societies. *Science* 312, 1767–1770.

Hoffman, E., McCabe, K., and Smith, V. (1996). On expectations and the monetary stakes in ultimatum games. *Intl. J. Game Theory.* 25, 289–301.

Insel, T.R. and Young, L.J. (2001). The neurobiology of attachment. *Nat. Rev. Neurosci.* 2, 129–136.

Kahneman, D., Knetsch, J.L., and Thaler, R. (1986). Fairness as a constraint on profit seeking: entitlements in the market. *Am. Econ. Rev.* 76, 728–741.

King-Casas, B., Tomlin, D., Anen, C. *et al*. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83.

Kirsch, P., Esslinger, C., Chen, Q. *et al*. (2005). Oxytocin modulates neural circuitry for social cognition and fear in humans. *J. Neurosci.* 25, 11489–11493.

Knoch, D., Pascual-Leone, A., Meyer, K. *et al*. (2006a). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314, 829–832.

Knoch, D., Gianotti, L.R.R., Pascual-Leone, A. *et al*. (2006b). Disruption of right prefrontal cortex by low-frequency repetitive transcranial magnetic stimulation induces risk-taking behavior. *J. Neurosci.* 26, 6469–6472.

Knoch, D., Nitsche, M.A., Fischbacher, U. *et al.* (2007). Studying the neurobiology of social interaction with transcranial direct current stimulation – the example of punishing unfairness. *Cerebral Cortex*, advance access (online).

Knutson, B., Adams, C.M., Fong, G.W., and Hommer, D. (2001a). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* 21, RC159.

Knutson, B., Fong, G.W., Adams, C.M. *et al*. (2001b). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport* 12, 3683–3687.

Knutson, B., Rick, S., Wimmer, G.E. *et al*. (2007). Neural predictors of purchases. *Neuron* 53, 147–156.

Koenigs, M., Young, L., Adolphs, R. *et al*. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature* 446, 908–911.

Kosfeld, M., Heinrichs, M., Zak, P.J. *et al*. (2005). Oxytocin increases trust in humans. *Nature* 435, 673–676.

Kringelbach, M.L. (2005). The human orbitofrontal cortex: linking reward to hedonic experience. *Nat. Rev. Neurosci.* 6, 691–702.

Ledyard, J. (1995). Public goods: a survey of experimental research. In: J. Kagel and A. Roth (eds), *Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press, pp. 111–194.

Levine, D.K. (1998). Modeling altruism and spitefulness in experiments. *Rev. Econ. Dynamics* 1, 593–622.

McCabe, K., Houser, D., Ryan, L. *et al*. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl. Acad. Sci. USA* 98, 11832–11835.

McClintock, C.G. and Liebrand, W.B.G. (1988). Role of interdependence structure, individual-value orientation, and another's strategy in social decision-making – a transformational analysis. *J. Pers. Social Psychol.* 55, 396–409.

Mikula, G. (1972). Reward allocation in dyads regarding varied performance ratio. *Zeitschrift Sozialpsychol.* 3, 126–133.

Moll, J., Zahn, R., de Oliveira-Souza, R. *et al*. (2005). Opinion: the neural basis of human moral cognition. *Nat. Rev. Neurosci.* 6, 799–809.

Moll, J., Krueger, F., Zahn, R. *et al*. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc. Natl. Acad. Sci. USA* 103, 15623–15628.

Montague, P.R. and Berns, G.S. (2002). Neural economics and the biological substrates of valuation. *Neuron* 36, 265–284.

Nitsche, M.A. and Paulus, W. (2001). Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans. *Neurology* 57, 1899–1901.

Nitsche, M.A., Schauenburg, A., Lang, N. *et al*. (2003). Facilitation of implicit motor learning by weak transcranial direct current stimulation of the primary motor cortex in the human. *J. Cogn. Neurosci.* 15, 619–626.

Nitsche, M.A., Doemkes, S., Karakose, T. *et al*. (2007). Shaping the effects of transcranial direct current stimulation of the human motor cortex. *J. Neurophysiol.* 97, 3109–3117.

O'Doherty, J.P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* 14, 769–776.

Poldrack, R.A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10, 59–63.

Preuschoff, K., Bossaerts, P., and Quartz, S.R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51, 381–390.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *Am. Econ. Rev.* 83, 1281–1302.

Ramnani, N. and Owen, A.M. (2004). Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nat. Rev. Neurosci.* 5, 184–194.

Rawls, J. (1972). *A Theory of Justice*. Oxford: Oxford University Press.

Rilling, J.K., Gutman, D.A., Zeh, T.R. *et al*. (2002). A neural basis for social cooperation. *Neuron* 35, 395–405.

Rilling, J.K., Sanfey, A.G., Aronson, J.A. *et al*. (2004). Opposing BOLD responses to reciprocated and unreciprocated altruism in putative reward pathways. *Neuroreport* 15, 2243–2539.

Rilling, J.K., Glenn, A.L., Jairam, M.R. *et al*. (2007). Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biol. Psych.* 61, 1260–1271.

Roth, A., Prasnikar, V., Okuno-Fujiwara, M., and Zamir, S. (1991). Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo: an experimental study. *Am. Econ. Rev.* 81, 1068–1095.

Samuelson, P.A. (1993). Altruism as a problem involving group versus individual selection in economics and biology. *Am. Econ. Rev.* 83(2), 143–148.

Sanfey, A.G., Rilling, J.K., Aronson, J.A. *et al*. (2003). The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1758.

Selten, R. (1998). Features of experimentally observed bounded rationality. *Eur. Econ. Rev.* 42, 413–436.

Sen, A. (1995). Moral codes and economic success. In: C.S. Britten and A. Hamlin (eds), *Market Capitalism and Moral Values*. Aldershot: Edward Elgar, pp. 23–34.

Seymour, B., Daw, N., Dayan, P. *et al*. (2007). Differential encoding of losses and gains in the human striatum. *J. Neurosci.* 27, 4826–4831.

Singer, T., Kiebel, S.J., Winston, J.S. *et al*. (2004). Brain responses to the acquired moral status of faces. *Neuron* 41, 653–662.

Singer, T., Seymour, B., O'Doherty, J.P. *et al*. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439, 466–469.

Slonim, R. and Roth, A.E. (1998). Learning in high stakes ultimatum games: an experiment in the Slovak Republic. *Econometrica* 66, 569–596.

Smith, A. (ed.) (1976). *The Theory of Moral Sentiments*. Oxford: Oxford University Press.

Smith, V.L. (1962). An experimental study of competitive market behavior. *J. Political Econ.* 70, 111–137.

Smith, V.L. (1982). Microeconomic Systems as an experimental science. *Am. Econ. Rev.* 72, 923–955.

Spitzer, M., Fischbacher, U., Herrnberger, B. *et al*. (2007). The neural signature of social norm compliance. *Neuron* 56, 185–196.

Thibaut, J.W. and Kelley, H.H. (1959). *The Social Psychology of Groups*. New York, NY: Wiley.

Tom, S.M., Fox, C.R., Trepel, C., and Poldrack, R.A. (2007). The neural basis of loss aversion in decision-making under risk. *Science* 315, 515–518.

Tricomi, E.M., Delgado, M.R., and Fiez, J.A. (2004). Modulation of caudate activity by action contingency. *Neuron* 41, 281–292.

van Lange, P.A.M. (1999). The pursuit of joint outcomes and equality in outcomes: an integrative model of social value orientation. *J. Pers. Soc. Psychol.* 77, 337–349.

van Lange, P.A.M., Otten, W., De Bruin, E.M.N., and Joireman, J.A. (1997). Development of prosocial, individualistic, and competitive orientations: theory and preliminary evidence. *J. Pers Social Psychol.* 73, 733–746.

van 't Wout, M., Kahn, R.S., Sanfey, A.G., and Aleman, A. (2005). Repetitive transcranial magnetic stimulation over the right dorsolateral prefrontal cortex affects strategic decision-making. *Neuroreport* 16, 1849–1852.

Veit, R., Flor, H., Erb, M. *et al*. (2002). Brain circuits involved in emotional learning in antisocial behavior and social phobia in humans. *Neurosci. Letts.* 328, 233–236.

Wagner, T., Valero-Cabre, A., and Pascual-Leone, A. (2007). Non-invasive human brain stimulation. *Annu. Rev.Biomed. eng.* 9, 527–565.

Wallace, B., Cesarini, D., Lichtenstein, P., and Johannesson, M. (2007). Heritability of ultimatum game responder behavior. *Proc. Natl. Acad. Sci. USA* 104, 15631–15634.

Winston, J.S., Strange, B.A., O'Doherty, J., and Dolan, R.J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nat. Neurosci.* 5, 277–283.

CHAPTER

16

The Study of Emotion in Neuroeconomics

Elizabeth A. Phelps

OUTLINE

INTRODUCTION

Neuroeconomic models of decision making have emphasized the assessment of value as a critical factor in driving choice behavior. Although the term *value* is used extensively in economic research, its precise definition can be somewhat elusive due to range of factors that might influence its determination (Bernoulli, 1954). Nevertheless, there is general agreement that the role of value is to provide a metric for the decision maker in weighing various options for action or choice. In the disciplines of psychology and neuroscience, another term that has been used to describe the appraisal or evaluation of events in the service of motivating action is *emotion* (Fridja, 1986; LeDoux, 1987; Scherer, 2005). Like value, emotion is a term whose precise definition can be elusive. Broadly speaking, emotion can be thought to be a "relevance detector" that lets us know what is important or significant (see, for example, Fridja, 1986). Under this broad category, however, it is widely recognized that the term *emotion*, as it is used in everyday language, represents a range of factors and processes that can be further delineated and assessed. Although a component of what might be considered emotion or affect overlaps with the concept of value, there are number of other emotion variables that are independent of value, but may also influence decision making.

By linking economic decision making to brain function, the emerging field of neuroeconomics has highlighted the overlap in the neural systems that mediate choice and other behaviors, including emotion. Like cognitive neuroscience before, the clean division between cognition (or reason) and emotion in economic decision making is blurred when attempting to understand the neural circuitry mediating these classes of behaviors (Phelps, 2006). Although emotion was considered an important variable in economic decision making prior to neuroeconomics (Mellers, 2000; Kahneman, 2003; Lowenstein and Lerner, 2003), the recent growth in this field has highlighted a role for emotion in economic choices (see, for example, Bechara *et al.*, 1997; Cohen, 2005; Shiv *et al.*, 2005). However, this growing interest in the role of emotion in decision making has rarely been coupled with the detailed investigation of the range of components, factors, and measures that have characterized the psychological study of emotion and affect.

The goal of this chapter is to introduce neuroeconomic researchers to some of the definitions, manipulations, and assessments of emotion that have been used in psychological and neuroscience research. To date, most neuroeconomic studies have depicted emotion as a single, unified variable that may drive choice, often in contrast to a reasoned analysis of the options and their relative values (see, for example, Cohen, 2005). This dual-system approach, although intuitively appealing, fails to consider the complexity of emotion or to capture the range of possible roles for emotion and affect variables in decision making. Adopting a more nuanced understanding of emotion will help clarify its impact on economic decision making and provide a basis for further understanding the complex interactions between emotion, value, and choice.

In affective science, there has been long-standing and considerable theoretical debate about the underlying structure of emotion and the parsing of emotional experience (for example, Barrett, 2006; Fridja, 2007a). These debates will not be reviewed here; rather, the focus will be on some of the more practical definitions, manipulations, and assessments that have advanced psychological research on emotion and affect, particularly as it relates to cognition and social behavior. When possible, how these different factors have been (or could be) used to advance economic and neuroeconomic research on decision making will be highlighted.

DEFINING COMPONENTS OF EMOTION AND AFFECT

One common approach to defining emotion has been to differentiate components of emotion and affect. Although there is debate about the properties of specific components, there is also significant overlap among component process theories of emotion (Fridja, 2007a). The framework presented below is largely based on the proposal of Scherer and colleagues (Scherer, 2000, 2005), with some modifications due to my own interpretation and emphasis, but many of the basic concepts have been suggested by other scholars as well (such as James, 1884; Cannon, 1929; Fridja; 1986, 2007b; LeDoux, 1987).

Emotion

Although the term *emotion* is commonly used to capture all affective experience, in component process theories of emotion and affect the term is proposed to reflect the discrete response to an external or internal event that entails a range of synchronized features, including subjective experience, expression, bodily response, and action tendencies. An additional important component of emotion is the evaluation and appraisal of the event. The characterization of emotion as a discrete, time-limited response to an internal or external event differentiates it from a range of other affective experiences. For an emotion to occur, it is not critical that all of the features are present, but rather that some subset are expressed in a relatively synchronized, temporally discrete manner. Each of these features is further characterized below.

Subjective Experience: Emotion vs Feeling

One of the primary sources of confusion in emotion research is the relation between emotion and feeling. The subjective experience of emotion, called *feeling*, is just one of the features that affective scientists consider a component of emotion. To laypersons reflecting on their experience, feelings are the consciously accessible and therefore most prominent characteristic of emotion, which may help explain why the terms emotion and feeling are often used interchangeably. In William James' seminal paper "What is Emotion"(1884), he differentiates feeling afraid from the perception and bodily response to seeing a bear, but early emotion theorists, including James, emphasized feelings as a critical characteristic of emotion. However, most emotion researchers today acknowledge that there are several responses that portray an emotion that do not depend on feeling (LeDoux, 1996). For research in non-human animals this distinction is critical, since subjective experience is not accessible in other species, but other types of emotional responses, such as physiological changes, are easily assessed and have characteristic patterns across species. In humans there is evidence the subjective experience of emotion does not always correspond with other expressions of emotion (Öhman *et al.*, 2000; Funayama *et al.*, 2001; Winkielman *et al.*, 2005), providing support for the suggestion that the subjective experience of emotion is neither critical nor necessary to determine that an emotion has occurred. Nevertheless a conscious, subjective experience, or feeling, is a common characteristic of emotion, as well as other affective states, such as mood (see below).

Expression

Both expression and bodily response might be considered reactions to emotion-eliciting events. They are distinguished here because of their potentially different roles in emotion. *Expression* refers here to motor responses in the face, voice or body that portray the emotion to others in a social environment. The expression of emotion has most often been studied in the characteristic motor response of the face when a person is experiencing an emotion. Darwin (1872/2002) suggested that the patterned facial expressions of emotion evolved for two functions, the first being as a means of social communication to allow conspecifics to both benefit from the emotional reactions of others, such as fear to a threatening stimulus or disgust to a noxious stimulus, and to determine the intent of others, such as smiling in appeasement or anger when threatening. The second proposed function of facial expressions, which has not been widely studied, is to alter the perceptual experience in adaptive ways by changing the facial configuration, such as widening the eyes to obtain more visual information in fear, or restricting the nasal passages to limit olfactory sensation in disgust.

The characteristic motor response for some basic facial expressions (see the section on Basic emotions, below) has been well characterized, and is thought to be universal and similar across cultures (Ekman and Friesen, 1971). There is substantial research into both the development of facial expression perception (e.g., Baenninger, 1994; Widen and Russell, 2003; Thomas *et al.*, 2007) and its underlying neural circuitry of facial expression perception (Adolphs, 2002). In contrast, relatively less is known concerning the vocal (Johnstone and Scherer, 2000) and bodily expression of emotion, although there is some evidence that the similar neural circuitry underlies the perception of emotional expressions across modalities (de Gelder, 2006).

Bodily Response

One of the unique characteristics of emotion is the patterned behavioral, hormonal, and autonomic responses that follow the perception of an emotion-eliciting event. Under this heading are also included neuromodulatory changes that characterize the impact of arousal and stress on the brain (LeDoux, 1996). In contrast to emotional expressions, in which a primary function might be the communication of emotion, bodily reactions are thought to be adaptive in preparing the organism to respond. A classic example is the "flight or fight" response, in which in the face of threat the sympathetic branch of the autonomic nervous system prepares the organism for quick action by changing the physiological state, such as increasing heart rate, blood pressure, respiration, and sweating (Cannon, 1929). In addition to hormonal, autonomic, and neuromodulatory responses there are also behavioral reactions – such as freezing, altering the amplitude of the startle reflex, or automatic withdrawal from a painful stimulus – that characterize the bodily response to emotion.

The characteristic patterns of expression and bodily response of emotion provide a powerful means to assess emotional reactions using psychophysiological techniques that are non-intrusive and do not depend on subjective experience or verbal report. In addition, the similarities in many of these patterns across species provide some assurance that studies of emotion in non-humans are capturing emotion reactions and qualities that are equally relevant to human experience.

Action Tendencies

In contrast to bodily responses, in which any behavioral motor actions that occur may be best characterized as automatic, reflexive reactions, emotion also elicits a tendency towards action that does not have a predictable motor pattern and is expressed as instrumental responses. These action tendencies motivate the organism towards a particular class of actions, such as approaching or withdrawing, but the actual action taken is modified by current goals and the situation. For instance, the tendency to move away from and avoid a stimulus that predicts potential threat is an instrumental response to an emotion-eliciting event. Exactly how the threat is avoided depends on the situation and options available. Similarly, there is a tendency to approach a stimulus that is rewarding, which can be expressed by a range of instrumental responses. The actual behavioral response, or action, is not determined by the emotion, but rather the emotional response motivates a class of potential actions, or tendencies.

Although action tendencies are assessed with instrumental responses that are constrained by the current goals and circumstances, and in this way they differ from automatic reactions, they are not necessarily under conscious control. For example, action tendencies can be expressed in a change in reaction time to emotion-eliciting events (e.g., Pratto and John, 1991), alterations in specific, motivationally congruent motor responses (e.g., Chen and Bargh, 1999); and the frequency of an instrumental response (Rescorla and Solomon, 1967), as well as choosing among options.

Evaluation and Appraisal

The primary function of emotion is to highlight the significance or importance of events so that these events receive priority in further processing. For an emotion to occur, there needs to be an assessment of the relevance of the internal or external emotion-eliciting event to the organism. Although there is some debate among affective scientists about the importance of different aspects of this assessment (Lazarus, 1984; Zajonc, 1984), there is general agreement that this assessment is a process with more than one relevant factor (Fridja, 2007b; LeDoux, 2007; Scherer, 2005). For the purposes of description two of these factors, evaluation and appraisal, are differentiated here, although these terms are often used interchangeably.

The *evaluation* of the relevance or significance of event can occur rapidly, without conscious awareness or cognitive interpretation (Zajonc, 1984). There is abundant psychological and neuroscience evidence indicating that the emotional significance of an event and some emotional reactions do not depend on being consciously aware of the event (LeDoux, 1996; Bargh and Chartrand, 1999). A well-known example of this fast, non-conscious evaluation of emotional events is the subcortical pathway for detecting the presence of auditory tones that have been paired with aversive shock in a fear conditioning paradigm. In a classic study, Romanski and LeDoux (1992) showed that rats demonstrate fear responses to a tone paired with shock, even when the auditory cortex is lesioned. These conditioned fear responses depend on subcortical pathways in which the auditory thalamus projects directly to amygdala, which mediates the expression of conditioned fear. The evidence of a fear response without sensory cortical processing is a strong demonstration that the evaluation of emotional significance can be rapid and independent of awareness. However, the presence of a subcortical pathway for eliciting emotion should not be taken to mean that cortical processing does not mediate most emotional responses, including those that are unconscious. There are many examples of emotional responses elicited by events that are not available to awareness, but could not be detected without cortical processing for perception and evaluation (e.g., Bargh and Chartrand, 1999; Whalen *et al*., 2004).

Although the emotional evaluation of an event can occur rapidly and without awareness, it is more often the case that we are aware of the emotional significance of an event. This awareness and the cognitive interpretation of the meaning of the event can initiate and alter an emotional response (Lazarus, 1984). The conscious assessment, cognitive monitoring, and interpretation of the significance of the event is referred to as *appraisal*. The appraisal of an event can occur rapidly, or can unfold as the circumstances surrounding the event provide additional cues to its significance. The importance of the appraisal process has been most often cited in relation to its impact on the subjective experience of emotion. In a classic study, Schacter and Singer (1962) evoked physiological changes consistent with emotion by administering epinephrine, but only informed a subset of participants that the bodily changes may be related to the drug. They then placed all the participants in the social circumstances that might evoke happiness or anger. The participants who were unaware of the cause of their bodily changes were more likely to report the subjective experience of euphoria or anger, and behave in a manner consistent with these emotions. In this case, the appraisal of the situation resulted in a change in the subjective experience and expression of emotion. These results, and others, suggest that the appraisal of an event may be closely tied to the subjective experience of emotion, but this does not imply that appraisal is linked solely to subjective experience. The appraisal of an event elicits and modifies all the features of emotion.

Mood

Emotions are characterized by synchronized, discrete responses to an event that may be high in intensity. In contrast, *mood* is a diffuse affect state characterized primarily by subjective feelings that are relatively enduring and of generally low intensity. Although a mood can be elicited by the appraisal of an event, it can also emerge with no apparent cause. The predominant feature of mood is the subjective experience, but moods can influence behavior in a number of ways that are generally less event-focused than emotion, attitudes, or preferences.

Attitude and Preference

Many affective scientists differentiate attitude and preference (e.g., Scherer, 2005), but they have been grouped here because they share many of the same properties. Like emotion, attitude and preference require the evaluation or appraisal of an event. Unlike emotion, attitudes and preferences are only elicited by external events such as another person, a group of people, an object, or a situation. The primary feature of attitudes and preferences is the affective evaluation of the external event. With *preferences* this evaluation is often described in terms of intrinsic liking or disliking, whereas with *attitudes* this evaluation is usually described in terms of valence of the affective response (i.e. good–bad). A key component of attitudes and preferences is that they result in a relatively enduring affective state linked to the properties or interpretation of the object or event. This affective state, while relatively enduring and consistent, can be modified somewhat by the situation, such as a change in preference for a particular food when sated. The affective state elicited by attitudes and preferences is generally described as low or medium intensity, and the primary behavioral impact is a tendency to approach or avoid the eliciting stimulus. Although attitudes and preferences both have subjective affective states and motivational components, social psychologists generally include a cognitive component (i.e. beliefs about the object) as a third important variable of attitudes (Breckler, 1984).

The notion of *value* in economics is most similar to the affective variable *preference*. Because of this, it is important carefully to consider the differences and similarities between preference and emotion in our efforts to understand their impact on economic decisions. As mentioned above, both preference and emotion are elicited by the evaluation and appraisal of external events, although emotions are also elicited by internal events. Preferences and emotions share two features; a subjective affective state and an action tendency. With preference, the subjective affective state is described as intrinsic liking or disliking. For emotion, it is described as feeling. Both preferences and emotions elicit action tendencies to approach or avoid that can be expressed with a range of behavioral responses, or not expressed at all.

In contrast to preference, emotion has two additional features: expression and bodily response. Emotion is also characterized as a synchronized, discrete response. Because of the time-limited nature of the response, emotion can be of higher intensity than might be expected with preference. With high-intensity emotions there may be a more immediate impact on behavior, or a faster disruption of ongoing behavior, but not all emotional reactions will have this quality. These different features suggest some of ways that the line between emotion and preference can become blurred. Although preferences are not described as eliciting expressions or bodily responses, it is certainly possible to like or dislike an object or event so much that this is expressed in the face or body. In addition, even though emotion is described as a discrete response, whereas a preference is more enduring, the same object can elicit both a consistent and discrete, synchronized response to its presentation, and a longer-lasting evaluation of liking or disliking. In short, the same stimulus can elicit both preference and emotion. The difference between the two may simply be a matter of degree. Given this, disentangling their impact in relation to the assessment of value and decision making may, at times, be difficult.

Other Affect Phenomena

The final two affect variables proposed by Scherer and colleagues (2000, 2005) are related to affective traits and situational style, and will only briefly be summarized here. *Affect dispositions* are relatively enduring traits that vary among individuals. For example, some individuals may be described as generally more happy, anxious, or angry than others. This persistent tendency over a range of situations and episodes to react in a relatively consistent manner suggests an affect disposition. In economic research, individual variability in factors such as sensitivity to risk across a range of situations would be similar to an affect disposition. Affect dispositions are also related to clinical disorders that have an affective component. *Interpersonal stance* is an affective style that colors interpersonal exchange in a social situation, and may be shaped by affect dispositions, attitudes and preferences, and strategic intentions. For example, a person with a tendency towards anger, when encountering someone she dislikes, may adopt a hostile approach to the social or economic exchange that may influence her actions and choices (see Scherer, 2005 for a review).

CATEGORIES OF EMOTION AND AFFECTIVE EXPERIENCE

When describing an affective experience, such as emotion, mood, or attitude, we are referring not only to the processes engaged but also to the quality of the state. For instance, it is rare that an emotion is mentioned without also labeling it using terms such as anger, happiness, guilt, etc. Below, two commons means of characterizing the quality of affective experience that have been used in psychological and neuroscience research are described. The approaches mentioned are used most frequently in affective research, but there are other proposed methods as well (see, for example, Scherer, 2005). It is important to note that these common categorization schemes may not always capture the range of affective experience that might be relevant to research on emotion and decision making. In general, the method of characterizing the affective state in studies of emotion is driven by the constraints of the experiment and means of assessment, as well as the specific question being addressed.

Basic Emotions

In his seminal work *The Expression of Emotion in Man and Animal*, Charles Darwin (1872/2002) proposed that there is a limited number of basic, universal human emotions. He derived this idea in part from colleagues who had studied different cultures around the world. When Darwin asked his colleagues about the emotional lives of individuals from other cultures, they reported similar emotional facial expressions. Darwin suggested that this universality of emotional expression suggests a common emotional experience.

More recently, Paul Ekman and his colleagues studied the facial expression of emotion and suggested that there are six basic emotional expressions: happy, sad, fear, anger, disgust, and surprise (Ekman and Friesen, 1971). Each of these expressions is characterized by a unique subset of facial muscle movements. The ability to convey these emotional expressions appears to be innate. Infants will display these facial expressions, as will individuals blind since birth who have not had the opportunity to view and mirror these expressions. Ekman has verified that these same facial expressions are observed across cultures, although the frequency of their expression is modified by cultural norms (Ekman, 1994). Studies examining the vocal expression of the emotion also provide some evidence for basic emotions (Johnstone and Scherer, 2000).

To the extent that facial expression indicates an emotional state, this research suggests that there may be basic, human emotions. Whether or not these patterned facial expressions are indicative of unique states has been a matter of debate (Barrett, 2006). Despite years of significant effort, there is relatively little evidence to suggest that these basic emotions are reflected in corresponding, unique patterns of autonomic responding (Cacioppo *et al.*, 2000a; Mauss *et al.*, 2005). Studies examining the neural circuitry mediating the perception of the basic emotions provide support for a unique neural circuitry for some basic emotional expressions, but there is also significant overlap between the neural patterns mediating the perception of different facial expressions (Calder and Young, 2005). In research on emotion, these basic facial expressions have proved useful in both assessing emotion perception and evoking corresponding emotional responses in others. However, it is important to acknowledge that these six basic emotions do not capture the range of human emotional experience. There are several more complex emotions, such as guilt and love, which are less clearly linked to specific facial or vocal displays.

Dimensions of Emotion

Another approach that has been used in the scientific investigation of emotion is to classify the range of emotion and affect states according to a few specific dimensions. The primary dimensional approach used in research on emotion and affect is captured by the *Circumplex* model. This model proposes that a number of emotional reactions and affective states can be characterized by two dimensions: arousal (activation–deactivation) and valence (pleasant–unpleasant) (Russell, 1980; Russell and Barrett, 1999). *Arousal* refers to the strength or intensity of the physiological response (or subjective judgment of physiological reaction) to a stimulus and the mobilization of energy. *Valence* reflects the degree to which the experience is pleasant (positive) or unpleasant (negative). Using these dimensions, the Circumplex model creates a framework for capturing a range of emotion and affect states. For example, sad, fearful, excited, and nervous are considered to be discrete emotional states. The Circumplex model suggests that these discrete states could be understood as varying along the dimensions of arousal and valence. Sad and fearful are both unpleasant, but sad is not as arousing or activating as fearful. Excited and nervous are both arousing states, but excited is relatively positive and nervous is relatively negative.

One advantage of this approach is that, by limiting the number of response categories to a few dimensions, the subjective judgment of the emotional state may be less influenced by the constraints of the question. In addition, the dimension of arousal is captured by physiological and neuromodulatory patterns that have been well characterized and can be assessed across species (Cannon, 1929; LeDoux, 1996). There are also some physiological indicators that capture valence (Lang *et al.*, 1990), although these may be most effective if the emotion-eliciting stimuli are also relatively high in arousal. Investigations of the neural systems mediating arousal and valence have shown some clear dissociations between regions sensitive to valence vs arousal (e.g., Anderson and Sobel, 2003), but there may be some brain regions most sensitive to specific arousal–valence combinations (see, for example, Cunningham *et al.*, 2008). One potential disadvantage of the using the Circumplex model to capture affective experience is that we rarely use dimensions of the emotion in everyday life. In discussing and describing emotion outside the laboratory, we tend to use more specific and nuanced descriptions of our affective state.

A second dimensional approach that is used less frequently by affective scientists but is particularly relevant to neuroeconomics is the approach/withdrawal distinction (Davidson *et al.*, 1990; Davidson, 2000). This approach classifies different emotions according to motivation. One of the primary functions of emotion is to motivate action, and different emotional states lead to different goals for action. Some emotional states, such as happiness, surprise, and anger, are referred to as *approach* emotions – that is, they evoke a motive or goal to approach a situation. Other emotional states, such as sadness, disgust, or fear, are *withdrawal* emotions, in that there is a natural tendency is to withdraw from situations linked to these emotions.

MANIPULATING AND MEASURING EMOTION IN RESEARCH

Our ability to investigate the impact of emotion on economic decisions and other behaviors is necessarily limited by our ability to manipulate and measure emotion in the laboratory. Theories of the underlying structure of emotion, such as the component process model outlined in this chapter, make several distinctions that may not always be fully captured by the techniques available. In spite of these limitations, psychological and neuroscience research has identified several techniques proven to be effective in understanding the impact of components of emotion on cognition that may be useful in investigating the impact of emotion and affect on economic decisions.

Manipulating Emotion

The manipulation of emotion and affect in the laboratory is constrained by ethical standards that prevent us from eliciting strong emotional reactions in humans for research purposes. For this reason, corroborating evidence from research with animal models and examples or studies from outside the laboratory may be particularly important. Below, three common techniques are outlined.

Emotion-eliciting Stimuli

Perhaps not surprisingly, the primary technique used to elicit emotion in the laboratory is to present emotionally evocative stimuli or situations. The question for emotion researchers is, what kinds of stimuli are most effective or appropriate to elicit the desired response? There are several classes of stimuli that have been well characterized and used across a range of studies.

Two classes of these stimuli are emotional scenes and words. Peter Lang and colleagues (1999) have developed a database with over a thousand complex scenes with varying emotional content, called the International Affective Picture System (IAPS). These scenes have been characterized by a large sample based on subjective ratings using the dimensions of valence and arousal. These norms obtained with subjective ratings have been confirmed for subsets of scenes with physiological measures that assess arousal and/or valence (Lang *et al.*, 1990, 1993). More recently, Lang and colleagues developed a similar set of emotion-eliciting words, (Affective Norms for English Words – ANEW) and sounds (International Affective Digital Sounds – IADS). These stimuli are made freely available to interested researchers, and have been widely used in studies of emotion and cognition (http://csea.phhp.ufl.edu/media.html).

Another class of stimuli commonly used in emotion research consists of faces with emotional expressions. As mentioned earlier, the muscle patterns of the six basic facial expressions have been extensively studied and characterized. The presentation of these faces has been shown to elicit a range of emotional responses in the perceiver, including those assessed with facial muscle movements (Tassinary and Cacioppo, 1992), subjective judgments (Zajonc, 1984; Adolphs, 2002), and choices (Winkielman *et al.*, 2005). Paul Ekman and colleagues have developed sets of pictures of facial expressions that have been used extensively in studies of emotion (http://www.paulekman.com/).

Other stimuli typically used to elicit emotion in laboratory studies would be considered *primary reinforcers* – that is, stimuli or situations that are inherently appetitive or aversive. For instance, mild electric shock is typically used in studies of conditioned fear (see, for example, LaBar *et al.*, 1998), and threat of shock is used to elicit an anxiety response (Grillon and Davis, 1997). To assure that electric shock is administered in an ethical manner, these studies generally use a work-up procedure in which participants indicate when the level of shock is "uncomfortable, but not painful." Other primary reinforcers used in research include juice for thirsty participants (Tobler *et al.*, 2006), and food for hungry participants (LaBar *et al.*, 2001; O'Doherty *et al.*, 2006). One advantage of these primary reinforcers is that they are also used in studies with non-human animals, allowing the investigation of emotional responses and neural systems using analogous tasks across species (see, for example, Phelps and LeDoux, 2005).

Another technique that might be considered a primary reinforcer in humans is to create social pressure or stress by presenting a social evaluation situation. In the Trier Social Stress test, research participants are asked to give a short public performance that will be evaluated. In some studies, participants are also asked to perform arithmetic in front of an audience. This brief social evaluation situation has been shown to elicit hormonal and physiological emotional responses (Kirschbaum *et al.*, 1993).

A final technique that has been used in emotion research that is highly relevant to neuroeconomic research is to give or take away money. Money, by definition, is a *secondary reinforcer*, in that its reinforcing properties are tied to what it represents; however its prominence as a *cultural reinforcer* may give it properties more analogous to a primary reinforcer than to a typical secondary reinforcer. Although money is the incentive used most frequently in economic studies of value and choice behavior, the presentation or removal of money has also been shown to elicit physiological responses indicative of an emotional reaction (Delgado *et al.*, 2006; Delgado *et al.*, 2008).

Mood Induction

Another technique that has been used to manipulate affect in the laboratory is focused on altering mood. This technique, called *mood induction*, attempts to change the baseline state reported by a research participant when arriving at the laboratory, and to have this change in mood persist throughout the experimental task. Typical means of changing a participant's mood are to present emotional film clips (e.g., scenes from a comedy routine to achieve a happy mood, or a death scene to achieve a sad mood), to play music that is depressing or upbeat, or to ask the participant to focus on emotional situations (real or imagined) that result in either positive or negative affect states, or a more specific state such as sadness or disgust (see, for example, Lerner *et al.*, 2004). Mood-induction procedures are considered successful if the participant reports a shift of mood state in the predicted direction using a subjective mood assessment.

Pharmacological Manipulations

A final technique that is not used frequently, but may be particularly useful in examining the neural basis of emotion, including neuroeconomic studies of emotion, is to introduce a drug that impacts emotion or social reactions. As described earlier, the classic study by Schacter and Singer (1962) mimicked a physiological arousal response by administering epinephrine. A pharmacological agent used in more recent human emotion research is propranolol, a beta-adrenergic antagonist that inhibits arousal responses in the body and brain. By administering propranolol, investigators have been able to further specify the impact of the physiological and neuromodulatory effects of arousal on memory and attention (Cahill *et al.*, 1994; DeMartino *et al.*, 2008).

A pharmacological agent that has been used previously in neuroeconomic research is oxytocin. Oxytocin is a hormone that is primarily known for its role in social bonding (Insel and Young, 2001). In a recent study, administering oxytocin was shown to increase social risk in a trust game (Kosfield *et al.*, 2005). Although oxytocin is primarily linked to social attachment, it has also been shown to impact emotional responses in social situations (Heinrichs, 2003). It has been suggested that the interaction of social cues and emotion may underlie some of the observed effects of oxytocin on behavior. A final pharmacological manipulation that may prove useful in future neuroeconomic and emotion research is to introduce dopamine agonists or antagonists, given the prominent role of dopamine in reward processing (see Chapter 21 of this volume). The advantage of pharmacological manipulations for neuroeconomic studies is that the impact of these drugs on the central nervous system is relatively well characterized by animal models. Because of this, they can provide unique insight into the neural mechanisms underlying emotion and decision making.

Measuring Emotion and Affect

Affective scientists have identified a range of techniques for assessing emotional experience. Most of these measure responses that are unique to emotion and affect, but some examine responses that could be indicative of range of behaviors, including emotion. One constraint of measuring emotion and affect in the laboratory is that it can be difficult to assess several types of emotional responses simultaneously, due to their complexity or interference between the measures. In addition, the assessment of emotion may, at times, influence the affective state being measured. Given this, researchers have to determine which measure is most appropriate for the specific question being addressed or emotion manipulation used.

Subjective Report

The primary means used to assess emotion and affect both in and outside the laboratory is to ask. Measures of subjective report often use Likert scales and ask participants to rate their emotional or affective state, or their reaction to an emotion-eliciting stimulus. The specific aspect of emotional experience assessed varies across studies, but often reflects the basic or dimensional approaches to categorizing emotional experience described earlier. One potential difficulty in using subjective reports of emotion is the possibility that asking participants to reflect on their affective state may alter their appraisal of the affective experience.

Physiological Measures

One of the unique aspects of emotion, in relation to other types of behaviors, is that there are several physiological responses accompanying emotional experience that can be assessed in a relatively non-intrusive manner. These physiological assessments provide a powerful means to examine emotion, both because they are unique to emotion and because they may represent specific autonomic and neuromodulatory response patterns that can provide additional insight into the underlying neural mechanisms. Here, two types of physiological responses used in emotion research are described; however, readers are referred to the *Handbook of Psychophysiology* (Cacioppo *et al.*, 2000b) for more detail, as well as information on additional physiological means for assessing emotion and affect.

The *skin conductance response* (SCR) is an indication of autonomic nervous system arousal. When someone becomes aroused, even if it is subtle, the sweat glands respond. This creates a change in the electrical conductivity of the skin. SCR is usually assessed by placing electrodes on the participant's fingers that pass a small electrical current through the skin. The participant does not feel anything from these electrodes. The electrodes pick up on subtle changes in the electrical conductivity of the skin with autonomic arousal. One advantage of SCR is that it is non-intrusive; participants do not have to do anything other than keep their fingers still. In addition, SCR can be assessed during functional magnetic resonance imaging (fMRI), allowing for concurrent assessment of arousal and blood oxygenation-level dependent (BOLD) signal. One potential disadvantages of SCR is that it takes a few seconds for an SCR in reaction to an event to emerge, so the presentation of stimuli has to be separated by several seconds. In addition, other responses, such a button press, can interfere with the assessment of SCR. SCR is a measure of arousal that does not differentiate positive or negative valence, or more specific categories of emotion experience.

A physiological measure that can be used to assess more specific emotion categories is *electromyography* (EMG). EMG is primarily used to assess the response of facial muscles in reaction to emotion-eliciting events. There are two primary ways EMG is used in the assessment of emotion. The first is as an index of the magnitude of the startle reflex. Startle is a reflex response that occurs when an individual is surprised, such as when hearing a sudden loud noise. The startle reflex is stronger or potentiated when experiencing a negative affective state, and reduced or attenuated somewhat during a positive affective state (Lang *et al.*, 1990). This can be measured in the laboratory by examining the strength of the eyeblink response to a loud, white-noise startle probe. One component of the startle reflex is an eyeblink, and the strength of this eyeblink response is measured by electrodes placed on the skin over the muscles around the eyes. For example, the startle reflex has been used as a physiological measure to assess the valence of the IAPS scenes described earlier. A brief, loud white noise was presented to participants as they viewed negative, positive, or neutral scenes, and in response to this loud white noise they blinked. The strength of this eyeblink, as measured by EMG, demonstrated an enhanced or potentiated startle reflex while viewing the negative scenes, relative to neutral scenes, and a slightly attenuated startle reflex while viewing positive scenes (Lang *et al.*, 1990). Some advantages of the startle response as a physiological measure of emotion are that it is a discrete response that can be assessed quickly, and it can indicate valence. Some disadvantages are that the loud, white-noise probe can be aversive to participants, and its measurement is difficult during fMRI, both because the startle reflex elicits head movement and because the electrodes are on the face inside the bore of the magnet, leading to significant interference in the EMG signal.

EMG has also been used to measure responses of muscles that indicate specific facial expressions. Given that the pattern of facial muscle movement differs for the six basic facial expressions described earlier, researchers have used EMG to assess facial movements consistent with different expressions. By measuring the activity of only a few specific facial muscles, researchers can reliably differentiate positive and negative emotional reactions. Although more specific facial expressions can be detected, they require the measurement of additional muscle movements, and these more complex response patterns are less frequently assessed. The assessment of facial muscle movement as an indication of an emotional reaction does not require that the participant make a facial expression that is apparent by visual inspection; even subtle movements of facial muscles that do not result in an easily observable facial expression can be detected (Tassinary and Cacioppo, 1992). Advantages of this technique are that it is non-intrusive and does not require any response by the participant, and it can detect relatively subtle emotional responses the measurement of which is quick and discrete. One disadvantage is that it may be difficult to assess concurrently with fMRI.

Other Assessment Techniques

Another non-intrusive technique to assess facial expressions in response to emotional events is visually to observe the face and code the patterns of muscle movements. Paul Ekman and colleagues have developed a Facial Action Coding System (FACS) that trains researchers to detect the response of specific muscle movements of the face by visual inspection. With training in FACS, the researcher can videotape the facial expression of research participants and code which basic facial expression is expressed during different stages of the experiment. This technique requires, of course, that the facial muscle movement be apparent to visual inspection, although subtle muscle movements that do not result in the full expression pattern can often be reliably observed.

Another common assessment of emotion uses reaction time. Studies have shown that an emotion-eliciting event can both increase and decrease reaction time, depending on the task. When emotion leads to increased reaction time, it is suggested that emotion interferes or conflicts with task processing (see, for example, Pratto and John, 1991; de Martino *et al.*, 2008). When emotion leads to a decrease in reaction time, it is suggested that emotion facilitates performance on the task (e.g. Öhman *et al.*, 2001). One of the primary difficulties in using reaction time as a measure of emotion is that it is non-specific. Reaction-time differences are generally used as measure of mental processing speed that can be indicative of a number of cognitive, behavioral, and emotional processes. For this reason, it is important that, when using reaction time as a measure of emotion, the task design does not vary additional factors (e.g, complexity, conflict) along with the emotion variable of interest. In addition, reaction time cannot indicate which component of emotion or affect may be linked to any observed difference in processing speed.

The measure of emotion or affect that is used in current economic studies of decision making is choice or action. The selection of one option over others is typically used as an indication of preference or attitude. The action taken in a task is used as an indication of an approach or withdrawal action tendency elicited by an emotion, mood, attitude, or preference.

EMOTION AND AFFECT IN ECONOMIC AND NEUROECONOMIC RESEARCH

Although there has been increasing interest in the interaction between emotion and decision making, there are surprisingly few economic and neuroeconomic studies that have explicitly manipulated or measured emotion or affect variables. In many studies and theories, emotion is inferred, but not directly altered or assessed. However, there have been some attempts to examine emotion variables in studies of economic choice. Here, for each component of emotion or affect described earlier, one economic or neuroeconomic study that has measured or manipulated this emotion or affect variable is highlighted. Finally, the problem of inferring emotion from patterns of brain activation in neuroeconomic research is discussed.

Emotion

As outlined earlier there are several features that characterize an emotion. One critical feature differentiating emotion from mood is that it is a discrete response to an event. Differentiating emotion from attitude and preference is a little more complex due to overlapping features in response to an event, such as subjective experience and action tendencies. For each feature of emotion, one study that has assessed or manipulated this variable is described briefly.

Subjective Experience

One anomaly in economic decision making is revealed in the ultimatum game, in which participants will reject a small monetary offer when the alternative is to receive nothing. A general interpretation of this effect is that when an offer is deemed to be unfair, the receiver will pay a cost to punish the offerer, who receives nothing if the offer is rejected. It is suggested this type of altruistic punishment may play a role in maintaining social norms (Fehr and Rockenbach, 2004).

In an effort to determine whether emotion plays a role in the decision to reject unfair offers in the ultimatum game, Pillutla and Murnighan (1996) asked participants to rate the offers received both in terms of fairness and their subjective experience of anger at receiving the offer. They found that different factors (such as type of knowledge about the game) differentially influenced ratings of fairness and anger, although they were correlated. In addition, they found that anger was a better explanation of rejections than the perception that the offers were unfair. Although a sense of fairness may be associated with an affective evaluation, by directly assessing an emotional response (subjective feelings), an additional factor that may mediate decisions in this economic game was identified. This result provides an example of the value of assessing emotion in studies of economic decision making.

Expression

There appear to be no studies that have assessed expression in the face, voice, or body as a measure of emotion in economic decision making. However, a recent study by Winkielman and colleagues (2005) used facial expressions as emotion-eliciting cues to examine the impact of emotion on both the consumption of a beverage and the subjective rating of preference. The goal was to demonstrate that participants do not need to be consciously aware of an emotional stimulus or response for it to influence choice behavior.

In this study, participants were presented pictures of faces with neutral, happy, or angry expressions. These faces were presented very briefly, and were immediately followed by a neutral face. Using these procedures, participants were unaware of the presentation of the faces with emotional expressions. Following the presentation of the faces, they were given the opportunity to pour and drink a beverage. They were also asked to rate their subjective feelings using a scale assessing valence and arousal. In a second experiment using a similar procedure, participants were asked to indicate how much they were willing to pay for the drink and whether they wanted more of the drink.

The study found that the observed effects were strongest for participants who rated themselves as thirsty. These participants poured and drank more, were willing to pay more, and were more likely to want more when presented subliminal happy faces compared to neutral faces. The opposite pattern was observed when participants were presented subliminal angry faces. Interestingly, the subliminal emotion manipulation did not influence subjective judgments of affect, providing support for the suggestion that emotion's influence on choice behavior can be independent of the subjective experience of emotion.

Bodily Response

One of the first and most influential neuroeconomics studies on emotion was the initial demonstration by Bechara and colleagues (Bechara *et al*., 1997) that physiological indicators of arousal are correlated with choices in an economic gambling task, and that damage to the orbitofrontal cortex impairs both the physiological response and the choice behavior.

The task used is known as the Iowa Gambling Task (IGT), and requires that participants pick a card from one of four decks over a series of trials. Each card selected represents a monetary reward or punishment, and the goal is to learn, through trial and error, which deck will yield a higher profit over time. The decks vary in their payoff schedules, with two of the decks (i.e. the "bad" decks) yielding greater rewards but also greater punishments, resulting in less overall profit than the other two decks (i.e. the "good" decks). Over a series of trials, normal, healthy participants eventually learned to select from the good decks more often than the bad decks, increasing their profits. Using SCR, physiological arousal was assessed as participants contemplated selecting from the good and bad decks. Higher arousal was observed prior to selecting from the bad decks. This enhanced arousal to the bad decks emerged prior to explicit knowledge about the reinforcement properties of the different decks. Bechara and colleagues (1997) suggested that it is this arousal response that steers participant away from the bad decks and promotes adaptive decisions.

A second group of participants with damage to the orbitofrontal cortex failed to learn to select more from the good decks over time. These patients also failed to generate anticipatory arousal responses. The results from this simple task are consistent with clinical reports of patients with damage to this region showing impaired social and emotional decision making in their daily lives. Based on these results, and others, Antonio Damasio (1994) proposed the Somatic Marker Hypothesis, which proposes that bodily states and emotional responses play a fundamental role in driving choice behavior.

In the years since the groundbreaking study by Bechara *et al*. (1997), a number of investigators have challenged the claims of the Somatic Marker

Hypothesis, and the interpretation of the IGT task in particular (see Dunn *et al.*, 2006 for a review). It has been proposed that the impairment observed with orbitofrontal cortex damage may be due to an inability to flexibly update stimulus–response contingencies, called reversal (Fellows and Farah, 2005). In addition, it has been suggested that explicit knowledge may underlie the observed pattern of choice behavior (Maia and McClelland, 2004), and that the relation arousal and choice is not specific (Tomb *et al.*, 2002). Finally, it has been shown that impairments in autonomic feedback do not always result in poor performance on the IGT (Heims *et al.*, 2004).

In spite of these concerns, the somatic marker hypothesis (Damasio, 1994) and the Bechara *et al.* (1997) study played a fundamental role in neuroeconomics and emotion. By assessing a physiological emotional response in an economic decision task and linking performance to human brain function, these findings highlighted the importance and value of considering brain function and emotion in efforts to understand economic decisions, at a time when these topics were rarely discussed.

Action Tendencies

The tendency to action is assessed by instrumental responses and choices, and it is in this way that most neuroeconomic studies assess this tendency. However, an especially powerful demonstration of an action tendency in emotion is the increased motivation to perform an instrumental response in the presence of an emotional cue that is unrelated to the instrumental response or choice. This change in motivation, sometimes referred to as *vigor*, is demonstrated in the Pavlovian-instrumental transfer paradigm (Rescorla and Soloman, 1967).

A recent neuroeconomic study examined the circuitry mediating Pavlovian-instrumental transfer in humans (Talmi *et al.*, 2008). Through simple pairing (Pavlovian conditioning), participants were trained to expect a monetary reinforcement when one stimulus was presented (the conditioned stimulus, CS+), but not with another (the CS−). This pairing led to faster reaction times and a higher pleasantness rating for the CS+. They were then trained to squeeze a handgrip to receive a monetary reward (instrumental conditioning). After partial extinction of the instrumental response, the Pavlovian-instrumental transfer test was conducted. In this final test stage the CS+ and CS− cues were presented at the same time as participants were given the opportunity to squeeze the handgrip. There was no monetary reinforcement in this final stage, so both the Pavlovian and instrumental responses were being extinguished. During this test phase, the participants squeezed the handgrip more frequently in the presence of the CS+ relative to the CS−. This increased frequency, or vigor, of the instrumental response in the presence of conditioned reward stimulus is indicative of an action tendency elicited by an emotion cue.

An examination of the pattern of BOLD response showed that activation of nucleus accumbens was related to the enhanced frequency of handgrip movements observed in the presence between the CS+. Across subjects, there was a correlation between the amygdala response and the magnitude of the Pavlovian-instrumental transfer effect. These results confirm studies from non-human animals suggesting that the action tendency elicited by conditioned emotional cues and expressed by instrumental responses may involve the interaction of the amygdala and striatum (Balliene, 2005). This type of enhanced motivation or action tendency in the presence of conditioned emotion cues is thought to play role in craving and drug-seeking (Everitt *et al.*, 2001), as well as economic behavior.

Evaluation and Appraisal

Changing how an event is appraised is fundamental to studies of emotion regulation. By interpreting an event differently, the emotional response to an event can be altered (Ochsner and Gross, 2005). A recent study adapted an emotion-regulation technique to examine the impact of altering appraisal on economic decisions and emotional reactions (Curley *et al.*, 2007). Participants were presented with a series of risky gambles and asked to choose between the gamble and a guaranteed outcome. For half of the gambles, participants were instructed to view each choice as one in a portfolio of choices. For the other half, they were instructed to focus on each choice as if it were the only choice.

The portfolio strategy resulted in an overall decrease in loss aversion. In addition, the more loss averse a participant was, the higher the physiological arousal response (as assessed with SCR) to monetary losses relative to gains at outcome. Those participants who were more successful at decreasing loss aversion with the portfolio strategy also showed a greater decrease in their arousal response to losses with the strategy. These results show that a typical economic behavior, loss aversion, may be linked to an emotional response, arousal, and that changing the appraisal of a choice may impact both of these responses.

Mood

Like emotion, mood can elicit action tendencies that may impact economic choices. In a demonstration of this, Lerner and colleagues (2004) induced moods of sadness and disgust and examined their impact on the endowment effect. Prior to the mood-induction procedure, half the participants were given a highlighter set for later use. Participants then watched one of three movies chosen to elicit a sad, disgust, or neutral mood. To enhance the mood-induction manipulation, participants were also asked to write about how they would feel if they were in the situation depicted in the film clip. After the mood-induction procedure, they were asked how much they would sell (if endowed) the highlighter set for, or how much they would choose to pay for it (if not endowed). They then rated their mood on a questionnaire assessing the intensity of a range of affective states.

The neutral mood condition yielded the classic endowment effect, that is participants would demand more to sell the product than they would choose to pay for it. The sad-mood group showed the opposite pattern; the sell price was lower than the choice price. Finally, the disgust-mood group showed no endowment effect; sell and choice prices were equivalent. These results suggest that mood can have a powerful effect on economic choices. In addition, the effect is specific to different mood states. Additional studies have demonstrated that an angry mood has a unique impact on risk and decision making (Lerner and Tiedens, 2006).

Attitudes and Preferences

Preferences are often assessed by the choice between options. This is how most neuroeconomic and economic studies assess an indication of preference. A recent neuroeconomic study also assessed the subjective judgments of preference and related these judgments to BOLD response patterns during a consumer purchasing task (Knutson *et al.*, 2007). In this study, subjects were presented with a series of consumer items and a price, and were asked to decide whether they were willing buy the item at that price. After scanning, subjects rated the items on desirability (i.e. preference) and indicated the price they would be willing to pay for the item. These measures of preference and the price differential between asking price and the willingness to pay were used to assess the BOLD response patterns underlying the impact of preference and price on consumer purchases.

The primary region that correlated with preference was the nucleus accumbens, consistent with a role for this region in representing value. Two regions, the medial prefrontal cortex and the insular cortex, were correlated with the price measure, specifically excessive prices. BOLD responses in all these regions predicted purchasing decisions above and beyond the self-report variables. This study demonstrates the value of combining several different sources of information, including subjective judgments of affect and brain-imaging results, to enhance the accuracy of predicting decisions.

Reverse Inference: Determining Emotion from BOLD Response Patterns

One trend that has emerged in neuroeconomic studies of emotion that is far less prevalent and accepted in affective science broadly is the inference of emotion from a pattern of BOLD signal response (e.g. Cohen, 2005). This technique, called *reverse inference*, has often been used in brain-imaging studies as a *post hoc* method to explain unanticipated results. However, the use of reverse inference as the primary technique for assessing a behavioral or mental function in a task is relatively new. In a recent paper, Russell Poldrack (2006) examined the assumptions underlying the use of reverse inference in brain-imaging studies. Given the prominence of this approach in neuroeconomic studies of emotion, it is important to understand these assumptions and the strengths and weaknesses of this approach, particularly as they pertain to neuroeconomic studies of emotion.

In brain imaging research, reverse inference refers to the practice of reasoning "backwards from the presence of brain activation to the engagement of a particular cognitive function" (Poldrack, 2006, p. 59). As Poldrack (2006) points out, reverse inference is a logical fallacy, but nevertheless has been shown to be useful in scientific investigation in the past. By conducting an analysis of the value of reverse inference in deducing a linguistic function from activation in a well-known language area (i.e., Broca's area), Poldrack (2006) concluded that reverse inference provides weak evidence for the engagement of a specific cognitive function. However, he also identifies three factors that may enhance confidence in this technique in brain imaging studies. To understand the value of reverse inference to studies of neuroeconomics, it would be useful to examine how these factors may be related to the inference of emotion from patterns of BOLD signal.

The first factor is the selectivity of response of a particular brain region. The more selective the response to the function in question, the higher the likelihood reverse inference provides useful information. To explore the notion of selectivity in brain structures commonly linked to emotion, I will briefly review the selectivity of two of these regions: the amygdala and the insular cortex. The amygdala is a small, almond-shaped structure in the medial temporal lobe that is best known for its critical role in fear learning (LeDoux, 1996). In addition, the amygdala has been implicated in a range of other emotion and non-emotion processes. The amygdala is thought to be important in the modulation of memory with arousal, the enhanced attention to threat stimuli, the perception of odor intensity, the perception of fear from faces, the perception of trust from faces, processing of other race faces, and the interpretation of social intent from inanimate objects (see Phelps, 2006 for a review). The amygdala has also been linked to the perception of biological motion (Bonda et al., 1996), the detection of oddball tones (Herry et al., 2007), the perception of sharp contours (Bar & Neta, 2007), and framing effects in economic decisions (De Martino et al., 2006). Although many, but not all, of these tasks could be characterized as having an emotional or social function, it is difficult to suggest a global theory that encompasses the range of tasks that may lead to activation of the amygdala precisely because there is little selectivity of the response (e.g., Whalen, 2007; Phelps, 2006; Cunningham, 2008). One reason for this lack of selectivity in brain imaging studies may be that the amygdala is composed of number of subnuclei with discrete functions (LeDoux, 2000), which are not differentiated by standard brain imaging techniques.

Activation of the insular cortex has been suggested to represent an emotion system in neuroeconomic studies (e.g., Sanfey et al., 2003). The insular cortex is a large region that lies within the lateral sulcus and is often divided into anterior and posterior regions. Like the amygdala, the insular cortex has been implicated in a range of emotion and non-emotion tasks. The role of the insular cortex that is probably most associated with emotion is the representation of somatic information, particularly as it relates to arousal and feelings (Damasio, 2000; Craig, 2003a; Critchley et al, 2004). However, the insular cortex has also been implicated in the perception of pain (Craig, 2003b), disgust (Phillips et al., 1997), empathy (Singer et al., 2004), gustatory processing (Small, 1999), conditioned taste aversion (Nerad et al., 1996), the resolution of interference in working memory (Nelson et al., 2003), object recognition (Bermudez-Rarroni et al., 2005), spatial learning (Nerad et al., 1996), pitch perception (Zatorre et al., 1994), and speech production (Ackerman & Riecker, 2003). Like the amygdala, many of the functions ascribed to the insular cortex have an emotion or social component, but others do not, and they are quite diverse. Perhaps not surprisingly there have been few attempts to propose a global theory of the insular cortex, although its role in the representation of specific functions has been characterized (e.g., Craig, 2003b). By identifying specific regions within the insular cortex, it may be possible to isolate more discrete functions, but the current literature does not suggest activation in this region is selective to emotion.

The second factor Poldrack (2006) proposes might add to the value of reverse inference in brain imaging is if the selectivity of the proposed function is enhanced by the activation of a number of regions that are working as a connected network. The notion that emotion may be processed by a unique network of interconnected brain regions is the principle underlying the limbic system, first proposed by MacLean in 1952 (see Kotter & Meyer, 1992; LeDoux, 1987). The original conceptualization of the limbic system identified a circuit of regions, including the hippocampus, amygdala and cingulate cortex, that may be specialized for processing emotion. Over the years as our understanding of neuroscience and emotion has progressed, the proposed limbic system has been modified and updated. However, in spite of extensive investigation, there has yet to be any defining criteria for regions that are, or are not, included as part the limbic system. In addition, there is little evidence that regions often included as part of the limbic system function as a network or system (Dagleish, 2004; LeDoux, 2000). Because of its historical prominence in theories of brain function, the term *limbic* is still frequently used today, but many affective scientists question whether the limbic system is still a useful concept (Kotter & Meyer, 1992; LeDoux, 2000). In fact, some have suggested that the use of term limbic and the corresponding implication that there is a clearly defined emotion system in the brain may actually impede scientific progress. This is driven by the observation that whether a task is designed to assess an emotional or cognitive function, and regardless of the corresponding behavioral assessment, when activation of a limbic area is observed it is often explained by reference to the fact that the limbic area mediates emotion. This concern was characterized by affective neuroscientist, Joseph LeDoux, who suggests:

> “We are, in other words, at a point where the limbic theory has become an off-the-shelf explanation of how the brain works. However, this explanation is grounded in tradition rather than data. Deference to the concept is inhibiting creative thought about how mental life is mediated by the brain.” (LeDoux, 2000, p. 159).

Given the complexity of behaviors encompassed by the term emotion, it is not surprising that there does not appear to a single system, limbic or otherwise, that mediates this broad range of functions. To advance our understanding of the role of emotion in neuroeconomic research, it might be necessary to abandon the notion of single emotion system, as many affective neuroscientists already have.

The third factor suggested by Poldrack (2006) that may enhance the value of reverse inference in studies of brain imaging is the size of the region of interest. The smaller the region, the more confident one might be about the inference, presumably because smaller regions may have more discrete functions. When emotion is deduced by reverse inference researchers may be referring to one of many regions, some of which, like the amygdala, are rather small, and others, like the insular cortex, are quite large. However, as mentioned earlier, in spite of its size, the function of the amygdala is quite diverse. For other smaller brain regions, for instance the nucleus accumbens, there may be more consistency in the proposed functions, which might provide additional assurance when inferring a mental construct from brain activation.

This brief analysis of the factors influencing the use of reverse inference in neuroeconomic studies of emotion suggests that this approach is no more effective or valid in studies of emotion than other research areas. For this reason, it is important when concluding a role for emotion in neuroeconomic research that there is support from both brain imaging and behavioral or physiological data (Poldrack, 2006). Although reverse inference is a powerful technique for generating hypotheses and ideas that inspire additional studies or measurements, its use as a primary technique for determining a role for emotion is questionable.

CONCLUSIONS

The distinction between emotion and cognition (or reason) has been prominent since early philosophical writings and this simple dichotomy continues to influence folk psychological theories and scientific thought, including economic and neuroeconomic research on decision making (Damasio, 1994). In neuroeconomics, this dual systems approach has gained prominence in studies attempting to characterize the impact of emotion on decision making (e.g., Cohen, 2005; Shiv, 2005). However, as research on the structure of cognition over last fifty years has demonstrated, cognition can best be construed as class of functions, including memory, language, attention, and reasoning, that may be interconnected, but also represent discrete processes that are independent. As outlined in this chapter, affective scientists similarly view emotion as representing a class of processes each of which is discrete, but also interconnected to other affective and cognitive functions. Research on the relation between emotion and cognition in cognitive neuroscience has slowly come to adopt the perspective that one cannot investigate the impact of emotion or affect on cognition without further specifying and assessing the specific emotion or affective process engaged (Phelps, 2006). Although this more nuanced approach to the study of emotion has influenced research on economic decision making (e.g., Lerner et al., 2004; 2006; Winkielman, 2005), its impact on neuroeconomics is less apparent. In this chapter, I have attempted to outline how neuroeconomics might benefit from adopting the tools of affective neuroscience in the study of emotion and decision making.

For economic and neuroeconomic researchers to understand the impact of emotion on decision making it may be important to further clarify the relation between the concepts of value and emotion. To the extent that emotion can be considered a "relevance detector" (Fridja, 1986; Scherer, 2005) that lets the organism know what is important, the overlap with the economist's view of value may be quite extensive. As outlined earlier, to the extent that value is viewed as an indication of liking or disliking, it may overlap with what emotion researchers call preference, which can be partially distinguished from other emotion and affect processes. Developing a more sophisticated approach to the study of affect in neuroeconomics will help define and specify the relation between value, emotion and choice.

References

Ackerman, H. and Riecker, A. (2003). The contribution of the insular to motor aspects of speech production: A review and a hypothesis. *Brain and Language* 89, 320–328.

Adolphs, R. (2002). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behav. Cogn. Neurosci. Rev.* 1, 21–62.

Anderson, A. and Sobel, N. (2003). Dissociating intensity from valence as sensory inputs to emotion. *Neuron* 39, 581–583.

Baenninger, M.A. (1994). The development of face recognition: featural or configurational processing? *J. Exp. Child Psychol.* 57, 337–396.

Balleine, B.W. (2005). Neural basis of food seeking: affect, arousal and reward in corticostriatal circuits. *Physiol. Behav.* 86, 717–730.

Bar, M. and Neta, M. (2007). Visual elements of subjective preference modulate amygdala activation. *Neuropsychologia* 45, 2191–2200.

Barrett, L.F. (2006). Are emotions natural kinds? *Persp. Psychol. Sci.* 1, 28–58.

Bargh, J.A. and Chartrand, T.L. (1999). The unbearable automaticity of being. *Am. Psychol.* 54, 462–479.

Bechara, A., Damasio, H., Tranel, D., and Damasio, A.R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science* 275, 1293–1295.

Bermudez-Rattoni, F., Okuda, S., Roozendaal, B., and mcGaugh, J.L. (2005). Insular cortex is involved in consolidation of object recognition memory. *Learning and Memory* 12, 447–449.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica* 22, 23–36.

Bonda, E., Petrides, M., Ostry, D., and Evans, A. (1996). Specific involvement of the human parietal systems and the amygdala in the perception of biological motion. *Jounal of Neuroscience* 16, 3734–3744.

Breckler, S.J. (1984). Empirical validation of affect, behavior, and cognition as distinct components of attitude. *J. Pers. Social Psychol.* 47, 1191–1205.

Cacioppo, J.T., Berntson, G.G., Larsen, J.T. *et al*. (2000a). The psychophysiology of emotion. In: M. Lewis and J.M. Haviland-Jones (eds), *Handbook of Emotions*, 2nd edn. New York, NY: Guilford Press, pp. 173–191.

Cacioppo, J.T., Tassinary, L.G., and Berntson, G.G. (eds) (2000b). *Handbook of Psychophysiology*. New York, NY: Cambridge University Press.

Cahill, L., Prins, B., Weber, M., and McGaugh, J.L. (1994). Beta-adrenergic activation and memory for emotional events. *Nature* 371, 702–704.

Calder, A.J. and Young, A.W. (2005). Understanding the recognition of facial identity and facial expression. *Nat. Rev. Neurosci.* 6, 641–651.

Cannon, W. (1929). *Bodily Changes in Pain, Hunger, Fear, and Rage*. New York, NY: Appleton.

Chen, M. and Bargh, J.A. (1999). Consequences of automatic evaluation: immediate behavioral predispositions to approach and avoid the stimulus. *Pers. Social Psychol. Bull.* 25, 215–224.

Cohen, J.D. (2005). The vulcanization of the human brain: a neural perspective on interactions between cognition and emotion. *J. Econ. Persp.* 19, 13–24.

Craig, A.D. (2003a). Interoception: The sense of the physiological condition of the body. *Current Opinion in Neurobiology* 13, 500–505.

Craig, A.D. (2003b). Pain mechanisms: Labeled lines versus convergence in central processing. *Annual Review of Neuroscience* 26, 1–30.

Critchley, H.D., Wiens, S., Rotshtein, P., Ohman, A., and Dolan, R.J. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience* 7, 189–192.

Cunningham, W.A., van Bavel, J.J., and Johnsen, I.R. (2008). Affective flexibility: evaluative processing goals shape amygdala activity. *Psychological Sci.* 19, 152–160.

Curley, N., Sokol-Hessner, P., Hsu, M. *et al*. (2007). Regulating loss aversion: a physiological correlate of the relative overvaluation of losses. *Proceedings of the Fifth Annual Meeting of the Society for Neuroeconomics*.

Dagleish, T. (2004). The emotional brain. *Nature Reviews Neuroscience* 5, 582–589.

Damasio, A.R. (1994). *Descartes Error: Emotion, Reason and the Human Brain*. New York, NY: G.P. Putnam & Sons.

Damasio, A.R. (2000). The Feeling of What Happen*s*. New York: Harcourt Brace & Company.

Darwin, C. (1872/2002). *The Expression of Emotion in Man and Animal*, 3rd edn. New York, NY: Oxford University Press.

Davidson, R.J. (2000). The neuroscience of affective style. In: R.D. Lane and L. Nadel (eds), *Cognitive Neuroscience of Emotion*. New York, NY: Oxford University Press, pp. 371–388.

Davidson, R.J., Ekman, P., Saron, C. *et al*. (1990). Approach/withdrawal and cerebral asymmetry: emotional expression and brain physiology. *J. Pers. Social Psychol.* 38, 330–341.

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nat. Rev. Neurosci.* 7, 242–249.

De Martino, B., Kumaran, D., Seymour, B., and Dolan, R.J. (2006). Frames, biases and rational decision-making in the brain. *Science* 313, 684–687.

Delgado, M.R., Labouliere, C.D., and Phelps, E.A. (2006). Fear of losing money? Aversive conditioning with secondary reinforcers. *Social Cogn. Affect. Neurosci.* 1, 250–259.

Delgado, M.R., Gillis, M.M., and Phelps, E.A. (2008). Regulating the expectation of reward via cognitive strategies. *Nature Neurosci.*, (in press).

de Martino, B., Strange, B.A., and Dolan, R.J. (2008). Noradrenergic neuromodulation of human attention for emotional and neutral stimuli. *Psychopharmacology* 197, 127–136.

Dunn, B.D., Dalgleish, T., and Lawrence, A.D. (2006). The somatic marker hypothesis: a critical evaluation. *Neurosci. Biobehav. Rev.* 30, 239–271.

Ekman, P. (1994). Facial expression and emotion. *Am. Psychol.* 48, 384–392.

Ekman, P. and Friesen, W. (1971). Constants across cultures in the face and emotion. *J. Pers. Social Psychol.* 17, 124–219.

Everitt, B.J., Dickinson, A., and Robbins, T.W. (2001). The neuropsychological basis of addictive behavior. *Brain Res. Rev.* 36, 129–138.

Fehr, E. and Rockenbach, B. (2004). Human altruism: economic, neural and evolutionary perspectives. *Curr. Opin. Neurobiol.* 14, 784–790.

Fellows, L.K. and Farah, M.J. (2005). Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cerebral Cortex* 15, 58–63.

Frijda, N.H. (1986). *The Emotions*. Cambridge: Cambridge University Press.

Frijda, N.H. (2007a). What might emotions be? Comments on the comments. *Social Sci. Inf.* 46, 433–443.

Frijda, N.H. (2007b). *The Laws of Emotion*. Mahwah, NJ: Lawrence Erlbaum.

Funayama, E.S., Grillon, C., Davis, M., and Phelps, E.A. (2001). A double dissociation in the affective modulation of startle in humans: effects of unilateral temporal lobectomy. *J. Cogn. Neurosci.* 13, 721–729.

Grillon, C. and Davis, M. (1997). Effects of stress and shock anticipation on prepulse inhibition of the startle reflex. *Psychophysiology* 34, 511–517.

Heims, H.C., Critchley, H.D., Dolan, R. *et al*. (2004). Social and motivational functioning is not critically dependent on feedback of autonomic responses: neuropsychological evidence from patients with pure autonomic failure. *Neuropsychologia* 42, 1979–1988.

Heinrichs, M., Baumgartner, T., Kirschbaum, C., and Ehlert, U. (2003). Social support and oxytocin interact to suppress cortisol and subjective responses to psychosocial stress. *Biol. Psych.* 15, 1389–1398.

Herry, C., Bach, D.R., Esposito, F., Di Salle, F., Perrig, W.J., Dcheffler, K., Luthi, A., and Seifritz, E. (2007). Processing of temporal unpredictability in human and animal amygdala. *Journal of Neuroscience* 27, 5958–5966.

Insel, T.R. and Young, L.J. (2001). The neurobiology of attachment. *Nat. Rev. Neurosci.* 2, 129–136.

James, W. (1884). What is an emotion? *Mind* 9, 188–205.

Johnstone, T. and Scherer, K.R. (2000). Vocal communication of emotion. In: M. Lewis and J.M. Haviland–Jones (eds), *Handbook of Emotions*, 2nd edn. New York, NY: Guilford Press, pp. 220–235.

Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *Am. Psychol.* 58, 697–720.

Kirschbaum, C., Pirke, K., and Hellhammer, D.H. (1993). The “Trier Social Stress Test” – a tool for investigating psychobiological responses in a laboratory setting. *Neuropsychobiology* 28, 76–81.

Knutson, B., Rick, S., Wimmer, G.E. *et al*. (2007). Neural predictors of purchases. *Neuron* 53, 147–156.

Kosfeld, M., Heinrichs, M., Zak, P.J. *et al*. (2005). Oxytocin increases trust in humans. *Nature* 435, 673–676.

Kotter, R. and Meyer, N. (1992). The limbic system: A review of its empirical foundation. *Behavioural Brain Research* 31, 105–127.

LaBar, K.S., Gatenby, J.C., Gore, J.C. *et al*. (1998). Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron* 20, 937–945.

LaBar, K.S., Gitelman, D.R., Parrish, T.B. *et al*. (2001). Hunger selectively modulates corticolimbic activation to food stimuli in humans. *Behav. Neurosci.* 115, 493–500.

Lang, P.J., Bradley, M.M., and Cuthbert, B.N. (1990). Emotion, attention and the startle reflex. *Psychological Rev.* 97, 377–395.

Lang, P.J., Greenwald, M.K., Bradley, M.M., and Hamm, A.O. (1993). Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology* 30, 261–273.

Lang, P.J., Bradley, M.M., and Cuthbert, B.N. (1999). *International Affective Picture System (IAPS): Technical Manual and Affective Ratings*. Gainesville, FL: University of Florida, Center for Research in Psychophysiology.

Lazarus, R.S. (1984). On the primacy of cognition. *Am. Psychol.* 39, 124–129.

LeDoux, J.E. (1987). Emotion. In: V.E. Mountcastle, F. Plum, and S.R. Geiger (eds), *Handbook of Physiology*, Vol. 5. Bethesda, MD: American Physiological Society, pp. 419–460.

LeDoux, J.E. (1996). *The Emotional Brain*. New York, NY: Simon and Schuster.

LeDoux, J.E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience* 23, 155–184.

LeDoux, J.E. (2007). Unconscious and conscious contributions to the emotional and cognitive aspects of emotions: a comment on Scherer’s view of what an emotion is. *Social Sci. Inf.* 46, 395–405.

Lerner, J.S. and Tiedens, L.Z. (2006). Portrait of the angry decision maker: how appraisal tendencies shape anger’s influence on cognition. *J. Behav. Decision Making* 19, 115–137.

Lerner, J.S., Small, D.A., and Loewenstein, G. (2004). Heart strings and purse strings: effects of emotions on economic transactions. *Psychological Sci.* 15, 337–341.

Loewenstein, G. and Lerner, J.S. (2003). The role of affect in decision making. In: R. Davidson, K. Scherer, and H. Goldstein (eds), *Handbook of Affective Science*. New York, NY: Oxford University Press, pp. 619–642.

Maia, T.V. and McClelland, J.L. (2004). A reexamination of the evidence for the somatic marker hypothesis: what participants really know in the Iowa gambling task. *Proc. Natl Acad. Sci. USA* 101, 16709–16710.

Mauss, I.B., Levenson, R.W., McCarter, L. *et al*. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion* 5, 175–190.

Mellers, B.A. (2000). Choice and the relative pleasure of consequences. *Psychological Bull.* 126, 910–924.

Nelson, J.K., Reuter-Lorenz, P.A., Sylvester, C.Y., Jonides, J., and Smith, E.E. (2003). Dissociable neural mechanisms underlying response-based and familiarity-based conflict in working memory. *Proceedings of the National Academy of Sciences* 100, 11171–11175.

Nerad, L., Ramirez-Amaya, V., Ornsby, C.E., and Bermudez-Rattoni, F. (1996). Differential effects of anterior and posterior insular cortex lesions on the acquisition of conditioned taste aversion and spatial learning. *Neurobiology of Learning and Memory* 66, 44–50.

Ochsner, K.N. and Gross, J.J. (2005). The cognitive control of emotion. *Trends Cogn. Sci.* 9, 242–249.

O’Doherty, J.P., Buchanan, T.W., Seymour, B., and Dolan, R.J. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron* 49, 157–166.

Öhman, A., Flykt, A., and Esteves, F. (2001). Emotion drives attention: detecting a snake in the grass. *J. Exp. Psychol. Gen.* 127, 69–82.

Phelps, E.A. (2006). Emotion and cognition: insights from studies of the human amygdala. *Annu. Rev. Psychol.* 57, 27–53.

Phelps, E.A. and LeDoux, J.E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* 48, 175–187.

Phillips, M.L., Young, A.W., Senior, C., Brammer, M., Andrew, C., Calder, A.J., Bullmore, E.T., Perrett, D.I., Rowland, D., Willaims, J.A., and David, A.S. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature* 389, 495–498.

Pillutla, M.M. and Murnighan, J.K. (1996). Unfairness, anger and spite: emotional rejections of ultimatum offers. *Org. Behav. Human Dec. Proc.* 68, 208–224.

Poldrack, R.A. (2006). Can cognitive process be inferred from neuroimaging data. *Trends Cogn. Sci.* 10, 59–63.

Pratto, F. and John, O. (1991). Automatic vigilance: the attention-grabbing power of negative social information. *J. Pers. Social Psychol.* 61, 380–391.

Rescorla, R.A. and Solomon, R.L. (1967). Two-process learning theory: relationships between Pavlovian conditioning and instrumental learning. *Psychological Rev.* 74, 151–182.

Romanski, L.M. and LeDoux, J.E. (1992). Equipotentiality of thalamo-amygdala and thalamo-cortico-amygdala circuits in auditory fear conditioning. *J. Neurosci.* 12, 4501–4509.

Russell, J.A. (1980). A circumplex model of affect. *J. Pers. Social Psychol.* 39, 1161–1178.

Russell, J.A. and Barrett, L.F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *J. Pers. Social Psychol.* 76, 805–819.

Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., and Cohen, J.D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1757.

Schacter, S. and Singer, J. (1962). Cognitive, social and physiological determinants of emotional state. *Psychological Rev.* 29, 379–399.

Scherer, K.R. (2000). Psychological models of emotion. In: J. Borod (ed.), *The Neuropsychology of Emotion*. Oxford: Oxford University Press, pp. 137–162.

Scherer, K.R. (2005). What are emotions? And how can they be measured? *Social Sci. Inf.* 44, 695–729.

Shiv, B., Loewenstein, G., Bechara, A. *et al*. (2005). Investment behavior and the negative side of emotion. *Psychological Sci.* 16, 435–439.

Singer, T., Seymour, B., O’Doherty, F., Kaube, H., Dolan, R.J., and Frith, C.D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science* 303, 1157–1162.

Talmi, D., Seymour, B., Dayan, P., and Dolan, R.J. (2008). Human Pavlovian-instrumental transfer. *J. Neurosci.* 28, 360–368.

Tassinary, L.G. and Cacioppo, J.T. (1992). Unobservable facial actions and emotion. *Psychological Sci.* 3, 28–33.

Thomas, L.A., de Bellis, M.D., Graham, R., and LaBar, K.S. (2007). Development of emotional facial recognition in late childhood and adolescence. *Dev. Sci.* 10, 547–558.

Tobler, P.N., O’Doherty, J.P., Dolan, R.J., and Schultz, W. (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *J. Neurophysiol.* 95, 301–310.

Tomb, I., Hauser, M., Deldin, P., and Caramazza, A. (2002). Do somatic markers mediate decisions on the gambling task? *Nat. Neurosci.* 5, 1103–1104.

Whalen, P.J., Kagan, J., Cook, R.G. *et al*. (2004). Human amygdala responsivity to masked fearful eye whites. *Science* 306, 2061.

Widen, S.C. and Russell, J.A. (2003). A closer look at preschoolers' freely produced labels to facial expressions. *Dev. Psychol.* 39, 114–128.

Winkielman, P., Berridge, K.C., and Wilbarger, J.L. (2005). Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Pers. Social Psychol. Bull.* 31, 121–135.

Zajonc, R.B. (1984). On the primacy of affect. *Am. Psychol.* 39, 117–123.

Zattore, R.J., Evans, A.C., and Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neurocience* 14, 1908–1919.

CHAPTER

17

Understanding Others: Brain Mechanisms of Theory of Mind and Empathy

Tania Singer

OUTLINE

INTRODUCTION

Social Neuroscience

In the past few years, the field of neuroscience has shown increased interest in the study of the affective and social brain. A new interdisciplinary field called *social neuroscience* has emerged from a union between classical cognitive neuroscience and social psychology. In recent decades, classical cognitive neuroscience has considerably advanced our understanding of how the brain processes shape, color, smells, and motion; recognizes objects; discriminates sounds; grasps actions; and enables us to perform higher-order cognitive operations like short- and long-term memory tasks, speech generation and recognition, and the executive functions involved in planning, multi-tasking, and self-monitoring. This approach, however, was governed by the implicit assumption that understanding a single brain is sufficient for understanding the behavior of all humans. Clearly, such an approach does not take into consideration the fact that humans are inherently social. Brains do not exist in isolation, and their basic functioning reflects their participation in the social culture into which they were born. It is likely that the differences between our brains and those of apes and monkeys are associated with our outstanding skills in social cognition, such as our ability to represent and understand the abstract beliefs and desires of others.

In general, social neuroscience seeks to understand phenomena in terms of the complex interactions between social factors and their influence on behavior, the cognitive processes underlying behavior, and finally the neural and hormonal mechanisms subserving cognitive processes (see also Ochsner and Lieberman, 2001). A multi-level and multidisciplinary approach such as this also requires the use of a multi-method research strategy, including methods as varied as behavioral measures (e.g., questionnaires, reaction times), neuroscientific imaging techniques (e.g., functional magnetic resonance imaging, fMRI, electroencephalogram, EEG; transcranial magnetic stimulation, TMS), and autonomic measures (e.g., heart rate, galvanic skin conductance).

At the beginning, social neuroscience focused predominantly on the investigation of basic social abilities (for a review and overview papers, see Adolphs, 1999, 2003; Ochsner and Lieberman, 2001; Blakemore *et al.*, 2004). Several functional imaging studies, for example, have investigated the neural correlates of attending, recognizing, and remembering socially relevant stimuli, such as the facial expressions of fear, attractiveness, trustworthiness, and racial identity, and the faces of fair and unfair players (Morris *et al.*, 1996; Hart *et al.*, 2000; Winston *et al.*, 2002; O'Doherty *et al.*, 2003; Singer *et al.*, 2004).

Another important line of research focuses on our ability to understand other people's minds – that is, their beliefs, intentions, and feelings. This line of research will be discussed in detail in the following sections of this chapter.

Yet another stream of research in social neuroscience has started to investigate moral and social reasoning in various ways. Moral reasoning is studied using *moral dilemma tasks*, which involve situations in which all possible solutions to a given problem are associated with undesirable outcomes (Greene *et al.*, 2001, 2004; Moll *et al.*, 2002a, 2002b, 2006; Greene, 2007; Moll and de Oliveira-Souza, 2007). *Social dilemma tasks* are closely related to but still distinct from moral dilemma tasks. Typically, social dilemma tasks involve strategies that differ with respect to the social desirability of their outcomes (e.g., self-serving vs cooperative). Social neuroscientists have used social dilemma tasks such as the simultaneous and sequential prisoners' dilemma game and the ultimatum game, which were developed in the framework of game theory, to investigate the neural underpinnings of social exchange and mutual cooperation. In studies employing these tasks, people play games for monetary payoffs using different playing strategies, some selfish and some cooperative, which allows for the investigation of social reasoning (figuring out what the other player will do; see, for example, Gallagher *et al.*, 2002; McCabe *et al.*, 2001; Rilling *et al.*, 2004), social emotions (emotional responses to fair and unfair play), and the interaction thereof (e.g., Montague *et al.*, 2002; Rilling *et al.*, 2002, 2007; Sanfey *et al.*, 2003; Singer *et al.*, 2004; King-Casas *et al.*, 2005). This is the stream of research that overlaps the most with or has even lead to the emergence of the new field of neuroeconomics.

Social Neuroscience and its Relation to Neuroeconomics and Decision Making

Even though the fields of social neuroscience and neuroeconomics are still perceived as two distinct fields, the topics they are concerned with overlap substantially, both in content and in methodology. Thus, researchers in both fields are interested in understanding the nature of human social interaction and human decision making and aim to determine the neural mechanisms underlying these complex social skills. Economic decision making, for example, frequently takes place in the context of social interactions. Game theory, developed in economics, has come to provide a very effective quantitative framework for studying how different pieces of information, incentives, and social knowledge influence strategies optimal for social interaction. In game theoretical paradigms, people typically engage in economic exchange tasks in the laboratory. One prominent example of a game that has also frequently been used in neuroscientific investigations is the ultimatum game (Sanfey *et al.*, 2003; Knoch *et al.*, 2006). In this game, Mover One (M1) is given a certain amount of money and can then decide how much she wants to share with Mover Two (M2). M2 looks at the offer and can then decide whether she wants to accept or reject it. If the offer is rejected, no one receives any money. Such a move can be conceived of as a way to punish M1. However, if M2 is purely interested in money, she should accept any possible offer from M1, irrespective of whether this offer is deemed fair or unfair. Another such game is the dictator game, in which M2 is neither allowed to reject nor accept the offers made by M1, but just passively receives whatever is offered.

But why does understanding how we understand others' minds matter in economic exchange? Let's go back to the above examples. When we compare the offers M1s typically make in ultimatum games to those made in dictator games, we find that M1s typically offer less in dictator games. The two games differ in that, in the ultimatum game, M2 is an active player who can influence the profits of both players. Thus, M1 has to construct a "*theory of mind*" of M2: What will M2 give me? Does M2 value fairness and, if so, which offer does M2 believe to be fair? How will M2 react if I give him X amount of money? More generally, the study of economic decision making in the context of game theory is based on the assumption that people can predict other people's actions when they understand their motivations, preferences, and beliefs (for a similar argument, see also Singer and Fehr, 2005; McCabe and Singer, 2008). However, economists still know little about the mechanisms that enable people to put themselves into other people's shoes and how these mechanisms interact with decision making in an economic context.

Social neuroscientists and neuroeconomists have recently started to clarify the neural mechanisms underlying our capacity to represent others' intentions, beliefs, and desires (referred to as *cognitive perspective-taking*, *theory of mind*, *mind-reading*, or *mentalizing*) and to share others' feelings (referred to as *empathy*). Even though our abilities to mentalize and to empathize are mostly used in concert when we try to understand other people's intentions, beliefs, desires, and feelings, preliminary evidence from studies of populations of patients with marked social deficits, like those with autism or psychopathy, suggest that mentalizing and empathizing are actually two distinct abilities that rely on distinct neural circuitries (see also Figure 17.1; Blair, 2005; Singer, 2006). For example, patients with autistic spectrum disorders often have deficits in cognitive perspective-taking, while psychopaths are very good at understanding other people's intentions and consequently at manipulating other people's behavior. In contrast, psychopaths lack empathy, which may be the reason for their antisocial behavior.

Such a dissociation points to an important difference between our abilities to mentalize and to empathize. Whereas both abilities play an important role in drawing inferences about other people's cognitive and emotional states, it has been suggested that empathy not only has an epistemological but also a motivational and social role (for a similar argument, see de Vignemont and Singer, 2006). Thus, empathy has very often been related to morality, altruism, justice, prosocial behavior, and cooperation (Batson and Shaw, 1991; Hoffman, 2000; Eisenberg and Morris, 2001). In the eighteenth century, the famous economist and philosopher Adam Smith (1759/2004: 1) alluded to the altruistic potential of human nature when he wrote about compassion or what he termed "fellow feeling" in *The Theory of Moral Sentiments*:

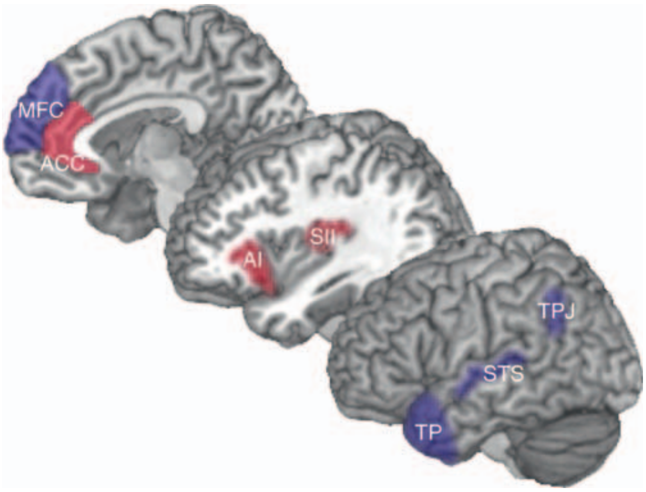


FIGURE 17.1 Brain networks involved in understanding others. Schematic representation of the brain areas typically involved in theory of mind (blue) and empathy (red) tasks. MPC, medial prefrontal cortex; ACC, anterior cingulate cortex; AI, anterior insula; SII, secondary somatosensory cortex; TP, temporal poles; STS, superior temporal sulcus; TPF, temporo-parietal junction.

> How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it. Of this kind is pity or compassion, the emotion which we feel for the misery of others, when we either see it, or are made to conceive it in a very lively manner. That we often derive sorrow from the sorrow of others, is a matter of fact too obvious to require any instances to prove it; for this sentiment, like all the other original passions of human nature, is by no means confined to the virtuous and humane, though they perhaps may feel it with the most exquisite sensibility. The greatest ruffian, the most hardened violator of the laws of society, is not altogether without it.

Accordingly, empathy is also likely to render people less selfish because it enables them to share others' emotions and feelings, which can help to motivate other-regarding behavior – in other words, behavior beneficial to another person and not only to oneself (e.g., helping someone).

Some behavioral evidence indeed suggests that people help others more when they report having empathized with them (Eisenberg and Morris, 2001). Empathy may also explain why people give charitable donations, an area of research which has also become the focus of neuroeconomic investigations (Chapter 20 of this volume; Moll *et al.*, 2006).

In the following, major findings concerning the neural mechanisms underlying our ability to understand the beliefs, intentions, motives, and feelings of other people will be summarized and discussed in light of possible implications for social and economic decision making.

DEFINING CONCEPTS

Clearly, the ability to understand other people's thinking and feeling is a fundamental component of our "social intelligence" and is needed for successful everyday social interaction. The literature very often refers to this as our capacity for *human empathy*. Even though, in lay terms, "empathy" usually refers to a unitary concept, a survey of the literature shows that empathy is a complex phenomenon composed of a variety of subskills and systems. It would go beyond the scope of this chapter to give a full account of existing definitions of empathy (for other relevant overviews from the fields of social neuroscience and psychology, see Wispe, 1986; Batson, 1987, 2008; Batson *et al.*, 1987; Eisenberg and Fabes, 1990; Hoffman, 2000; Preston and de Waal, 2002; Decety and Jackson, 2004; Keysers and Gazzola, 2006, 2007; Decety and Lamm, 2007). In this chapter, a neuroscientific perspective is taken, according to which three main systems rely on partially separable neural circuitries that all subserve our capacity to understand other people: (1) our ability to understand other people's motor intentions and action goals; (2) our ability to understand other people's beliefs and thoughts, which has been referred to as *theory of mind* (e.g., Premack and Woodruff, 1978), *mentalizing* (Frith and Frith, 2003), *mind-reading* (Baron-Cohen, 1995), or *cognitive perspective-taking*; and (3) our ability to understand other people's feelings, which is referred to as *empathy* or *emotional perspective-taking* (for a similar distinction, see Preston and de Waal, 2002; Gallese, 2003; Blair, 2005; de Vignemont and Singer, 2006; Singer, 2006; Decety and Lamm, 2007; Keysers and Gazzola, 2007).

De Vignemont and Singer (2006) define empathy as follows: we "empathize" with others when we have (a) an affective state, (b) which is isomorphic to another person's affective state, (c) which was elicited by observing or imagining another person's affective state, and (d) when we know that the other person's affective state is the source of our own affective state.

The first statement is important because it differentiates empathy from theory of mind, cognitive perspective taking, and mentalizing. The term "mentalizing" connotes a person's ability to cognitively represent the mental states of others, including their affective states, without becoming emotionally involved. The term "empathizing" connotes the capacity to share other people's feelings. Accordingly, when one "empathizes" with another person who is in pain, one feels the other person's pain in one's own body. In contrast, when one understands someone else's thoughts, one does not feel the thought of the other in one's own body. There are no qualia attached to it. This difference may become clearer when we consider the characteristics of psychopaths: Whereas psychopaths do not have an impaired ability to understand other people's wishes, beliefs, intentions, and desires, it appears that they do lack the embodied feeling of empathy, a feeling which enables those without psychopathy to anticipate others' suffering, thereby often preventing them from harming others. Thus, although psychopaths do possess the ability to mentalize, they are not able to empathize (for a similar argument, see Blair, 2005).

The second statement is important for the differentiation between empathy and *sympathy* or *compassion*. In all three cases, we feel vicariously for another person. However, when we empathize, we share another person's feelings; when we sympathize or show compassion we do not necessarily share the same feeling with the other person. For example, when I empathize with a person who feels sad, I feel sad myself, whereas when I sympathize or feel compassion for a person who feels sad, I feel either pity or love for the person but not necessarily sad. Also, when I notice that someone is jealous of me, I can sympathize with or show compassion towards that person, but I do not feel jealous myself. Further, when one empathizes with another person, there does not have to be a prosocial motivation attached to it – that is, a wish to maximize the other person's happiness or alleviate the other person's distress; when we sympathize or show compassion for another person, there is. For example, a torturer may use empathy in order to sense how to increase his victim's suffering, thus showing no compassion for his victim. Experiencing too much empathy can also lead to selfish instead of other-regarding behavior. For example, sharing too much of the other person's distress can lead to a withdrawal from – instead of helping – the suffering person. In general, however, empathy is conceived to be a first necessary step in a chain that begins with affect sharing, a subsequent understanding of another person's feelings, which then motivates other-related concern and finally engagement in helping behavior. Empathy and prosocial decision making are thus closely linked.

Finally, the last statement of the above definition of empathy is important for the differentiation between empathy and *emotional contagion*, whereby the latter connotes a reaction in which one shares an emotion with another person without realizing that the other person's emotion was the trigger. For example, babies start crying when they hear other babies crying, long before they develop a sense of a self separate from others. Recently, Neil Harrison and colleagues (2006) found initial evidence for pupillary contagion. When subjects were presented with photos of sad faces with different pupil sizes, their own pupil size mirrored that shown in the photos. Here, emotional contagion engaged the Edinger-Westphal nucleus in the brainstem which controls pupil size. Phenomena such as pupillary contagion occur involuntarily, and may represent a precursor of empathy. However, they are not considered "empathic responses," because the subjects are not aware that they are vicariously feeling for another person.

THE STUDY OF "THEORY OF MIND"

The History of Theory-of-mind Research

In 1978, Premack and Woodruff published a seminal paper in which they coined the term "theory of mind" (ToM) while discussing whether chimpanzees are capable of representing other primate's minds in terms of their desires, intentions, and beliefs. Despite extensive research conducted on this question after their paper was published, the debate about whether the capacity to have a theory of mind is uniquely human still has not been settled (Call, 2007). Overall, the literature appears to suggest that this ability is absent in monkeys, and only exists in a limited form in apes (Povinelli and Bering, 2002).

Around the same time, developmental psychologists also showed great interest in the study of the developmental time-course of our capacity to mentalize (for a review, see Frith and Frith, 2003). On the basis of a proposition by the philosopher Daniel Dennett (1978), who suggested that the most stringent test for the presence of ToM would be to see whether someone is able to predict someone else's actions on the basis of that person's false belief, Wimmer and Perner (1983) developed the false-belief paradigm to test children's mentalizing abilities. In the false-belief task, the following story is told: Maxi has some chocolate and puts it into a blue cupboard. Maxi leaves the room. Then his mother comes in and moves the chocolate to a green cupboard. Maxi comes back to get his chocolate. Where will Maxi look for the chocolate? A child who states that Maxi will look in the blue cupboard knows that he falsely believes the chocolate to be there. Control questions are posed to test whether the child understood the sequence of events: Where is the chocolate really? Do you remember where Maxi put the chocolate in the beginning? Another task, which is also used frequently in the field of ToM research, is the *Sally-Anne task*, in which Sally puts a ball in a basket and then Anne takes the ball out of the basket while Sally is out of the room. A series of studies using either of these tasks showed that children age four and older start to correctly attribute false beliefs to others and give verbal explanations when asked. At age five, over 90% of children understand this task, and at age six all of them do so (Baron-Cohen *et al.*, 1985; Perner *et al.*, 1987; for a review, see Frith and Frith, 2003). When the task is simplified with a little game and does not use verbal report as a dependent measure, even children as young as three years of age seem to have an understanding of false beliefs (Clements and Perner, 1994). Research in the domain of autistic spectrum disorders suggests that the ability to mentalize is severely delayed in autism. The lack of a theory of mind in most autistic children could explain their observed failure in communication and social interaction (for a review, see Frith, 2001).

The Neural Foundation of Theory of Mind

With the development of modern imaging techniques, the study of our capacity to reason about other people's minds has become the focus of cognitive neuroscience research. Imaging studies performed with healthy adults have used different paradigms to investigate which neural structures underlie our capacity to reason about other people's non-observable internal states. In these studies stories are typically told, on the basis of texts, abstract moving shapes, or cartoons, to subjects in the scanner, who are asked to understand the intentions, beliefs, and desires of the protagonist in the respective stories (for a review, see Gallagher and Frith, 2003). Theory-of-mind studies have consistently shown the involvement of a network comprising the posterior superior temporal sulcus (STS) extending into the temporoparietal junctions (TPJ), the medial prefrontal cortex (mPFC), and sometimes also the temporal poles (TP). A schematic representation of the mentalizing brain network is illustrated in Figure 17.1 in blue.

Frith and Frith (1999) suggested that the mPFC may represent mental states decoupled from reality, while the STS helps process a causal relationship between visual motion/action and another person's intended goals, and the temporal poles draw on encodings of past experience to "simulate" another person's experience. Recently, Rebecca Saxe suggested that different subcomponents of ToM have different developmental time-courses and rely on different brain regions. In line with earlier approaches in developmental psychology and philosophy, she proposed that the ability to understand mental state concepts like desires, goals, and feelings develops earlier than the ability to represent the more abstract contents of mental states, such as beliefs, and that the former relies on functions of the mPFC, whereas the latter is specifically associated with TPJ functions (Saxe and Wexler, 2005; Saxe and Powell, 2006).

Game theoretical paradigms have also been used to investigate mentalizing (Gallagher *et al.*, 2002; McCabe *et al.*, 2001; Rilling *et al.*, 2004). Subjects are scanned while playing strategy games against someone sitting outside the scanning room. For example, Gallagher *et al.* (2002) and McCabe *et al.* (2001) compared the brain areas involved when subjects played against another person with those involved when subjects played against a computer. These studies have repeatedly demonstrated medial prefrontal lobe involvement.

The mPFC is not only involved when people mentalize about other people's thoughts, intentions, and beliefs, but also when people are reflecting on their own states (Mitchell *et al.*, 2005a). Jason Mitchell (Mitchell *et al.*, 2002, 2005b, 2006) recently conducted a series of interesting studies on mentalizing that suggest that there are functional differences between judging the mental states of similar and dissimilar others. A more ventral part of the mPFC was recruited when participants made self-judgments or judgments about people whom they perceived as being similar to themselves with respect to appearance or political attitudes. By contrast, a more dorsal part of the mPFC showed enhanced activation – close to the activation found in the mentalizing studies cited above – when subjects judged the mental states of people perceived as being dissimilar to themselves. This suggests that we may use two different strategies when inferring other people's mental states: With one strategy we simulate the other person on the basis of knowledge we have about ourselves; with the other strategy we infer the mental states of the other person on the basis of more abstract knowledge we have acquired about the world. The latter strategy may also involve knowledge about stereotypes, and raises the interesting question as to whether judging another person's mental state may be biased in different ways depending on whether we perceive them as similar or dissimilar to ourselves. *Egocentric bias*, the propensity to understand other people's states in terms of our own, may easily occur if we simulate others on the basis of ourselves while ignoring possible differences between ourselves and others. In addition, misattributions may occur when we judge other people's mental states on the basis of stereotyped or categorical knowledge that underestimates the similarity between the other person and ourselves. An interesting twist for future research in this domain would be to explore whether brain data can predict what decisions are made during social exchanges based on whether ventral or dorsal mPFC is activated, which would reflect an unconscious perception of people being similar or dissimilar to ourselves.

THE STUDY OF EMPATHY AND FEELINGS

The Mirror Neuron System and its Role for Action Understanding

While theory-of-mind research focuses on complex inferences about abstract mental states such as another person's beliefs, another line of neuroscientific endeavor has focused on our ability to understand other people's goals and intentions by merely observing their actions. This line of research originated with the seminal discovery by a group in Parma, Italy, that neurons in the premotor cortex of macaque monkey brains were firing both when a monkey performed hand movements itself and when it merely observed another monkey or a human performing the same hand movements (Gallese *et al.*, 1996; Rizzolatti *et al.*, 1996; Ferrari *et al.*, 2003). These so-called *mirror neurons* were the first evidence for a brain mechanism which not only represents the subject's own world, but also another person's. It was suggested that mirror neurons may represent the basis for imitation. Thus, when imitating someone else's actions, we first have to transform what we see (action perception) into our own motor programs which allow us to generate a certain action sequence. The discovery of mirror neurons suggested the existence of such translation mechanisms in the primate brain.

Since the discovery of mirror neurons, several studies have demonstrated a similar common coding of the perception and generation of motor actions in the human brain using imaging techniques like PET and fMRI (for a review, see Grezes and Decety, 2001). In these studies, people were scanned while they watched movies depicting short motor actions. The observed activation was then compared to that observed when the scanned subjects performed the same motor action themselves. In line with the studies on monkeys, these studies revealed that the same circuitry was recruited when subjects merely observed another person performing an action and when they performed the same action themselves. These shared neuronal representations included brain regions in supplementary motor area (SMA), pre-SMA, premotor cortex, the supramarginal gyrus, intraparietal sulcus, and superior parietal lobe.

Currently, researchers are debating about the exact function of the mirror neuron system and its role in social cognition. Some authors have suggested that the mirror neuron system might play a general role in understanding other people's intentions and goals by providing us with an automatic simulation of their actions (Gallese and Goldman, 1998; Fogassi *et al.*, 2005). This suggestion is in line with simulation theoretical accounts developed in philosophy in the context of theory-of-mind research (Gordon, 1986; Heal, 1986; Harris, 1989). Simulation theory holds that what lies at the root of our mature mind-reading abilities is the ability to project ourselves imaginarily into another person's perspective by *simulating* their mental activity using our own (for an overview, see Carruthers and Smith, 1996). Simulation theorists originally opposed the so-called "theory theorists" who suggest that we understand other people's minds because we have acquired abstract knowledge and a folk psychological theory about the world and about the structure and functions of our own and other people's minds (Wellman, 1990; Gopnik and Wellman, 1994). Simulation approaches have now been extended to the domain of actions and feelings. To understand what another person is doing, we simulate their movements using our own motor program; to understand what other people are feeling, we simulate their feelings using our own affective programs (see also de Vignemont and Singer, 2006; Keysers and Gazzola, 2006). Whereas these accounts see the specific role of mirror neurons in the understanding of others' motor actions and action-related intentions, but not their feelings, Iacoboni and colleagues have suggested a motor theory of empathy according to which mirror neurons have a much broader role in social cognition in that their role is not only to understand action but also to understand others' emotions and minds (Carr *et al.*, 2003; Dapretto *et al.*, 2006; Iacoboni and Dapretto, 2006). Others have criticized this broad interpretation of the role of mirror systems, believing that it overemphasizes the role of motor mirror neurons in social cognition. They maintain that mirror neurons may help us to understand other people's simple, observable action goals, but not their abstract beliefs as usually conceptualized in theory-of-mind tasks (Jacob and Jeannerod, 2005; Saxe, 2005).

Empathy: A Shared Network Hypothesis

The idea that perception–action links in our brains enable us to understand others has recently been expanded to include the ability to share not only motor actions but also feelings and sensations with others (Preston and de Waal, 2002; Gallese, 2003; Decety and Jackson, 2004; Decety and Lamm, 2006; de Vignemont and Singer, 2006). Thus, in addition to the ability to understand action intentions or more abstract mental states such as other people's beliefs or wishes, humans can also empathize with others – that is, share and understand feelings and emotions. Humans can feel empathy for other people in a wide variety of contexts – when others feel basic primary emotions and sensations such as anger, fear, sadness, joy, pain, and lust, as well as more culturally variable secondary emotions like embarrassment and shame. Inspired by earlier perception–action models (Prinz, 1990) in the domain of action understanding, Preston and de Waal (2002) proposed a neuroscientific model of empathy suggesting that observing or imagining another person in a particular *emotional* state automatically activates a representation of that state in the observer, with its associated autonomic and somatic responses. The term "automatic" in this case refers to a process that does not require conscious and effortful processing, but which can nevertheless be inhibited or controlled.

Indeed, fMRI studies in humans have provided evidence for a role of such shared neural networks that enable one to feel – by merely perceiving or imagining another person feeling pain, touch, or disgust in the absence of any stimulation to one's own body – what it feels like for the other person to be in pain, touched, or disgusted. For example, some studies have been able to demonstrate that similar neural responses in anterior insula cortex (see Figure 17.1) – a brain region involved in processing, among other sensations, disgust and taste – are elicited when subjects view pictures of disgusted faces and when they smell disgusting odors themselves (Wicker *et al.*, 2003), or when subjects view videos showing people sampling pleasant or unpleasant tastes and when they sample the different tastes themselves (Jabbi *et al.*, 2007). In contrast, another study found shared activation in secondary somatosensory cortices when subjects watched videos of people being touched and when they were being touched themselves (again, see Figure 17.1). These results are in line with the role of somatosensory cortices for the processing of touch (Keysers *et al.*, 2004).

The majority of studies on empathic brain responses have been conducted in the domain of pain (Morrison *et al.*, 2004, 2007; Singer *et al.*, 2004, 2006; Avenanti *et al.*, 2005, 2006; Botvinick *et al.*, 2005; Jackson *et al.*, 2005, 2006; Bufalari *et al.*, 2007; Cheng *et al.*, 2007; Gu and Han, 2007; Lamm *et al.*, 2007;

Moriguchi *et al.*, 2007; Morrison and Downing, 2007; Saarela *et al.*, 2007). For example, in an early study, Singer and colleagues (2004) recruited couples and measured empathy *in vivo* by assessing brain activity in the female partner while painful stimulation was applied either to her own or to her partner's right hand via electrodes attached to the back of the hand. The male partner was seated next to the MRI scanner, and a mirror system allowed the female partner to see her own as well as her partner's hand lying on a tilted board in front of her. Before the experiment started, the couples were allowed to engage in social interaction to increase the feeling of being in a "real-life" situation. Differently colored flashes of light on a screen behind the board pointed to either the male or the female partner's hand, indicating which of them would receive painful and which would receive non-painful stimulation. This procedure enabled measurement of pain-related brain activation when pain was applied to the scanned subject (felt pain) or to her partner (empathy for pain). The results suggest that parts of the so-called *pain matrix* – the bilateral anterior insula (AI), rostral anterior cingulate cortex (ACC), brainstem, and cerebellum – were activated when subjects experienced pain themselves, as well as when they saw a signal indicating that a loved one had experienced pain. These areas are involved in the processing of the affective component of pain – that is, how unpleasant the subjectively felt pain is. Thus, both the experience of pain to oneself and the knowledge that a beloved partner is experiencing pain activates the same affective pain circuits, suggesting that if a beloved partner suffers pain, our brains also make us suffer from this pain. Activation in this network was also observed when subjects saw an unknown but likeable person suffering pain (Singer *et al.*, 2006), when subjects watched videos showing body parts in potentially painful situations (Jackson *et al.*, 2005, 2006), painful facial expressions (Lamm *et al.*, 2007), or hands being pricked by needles (Morrison *et al.*, 2004; for a review, see de Vignemont and Singer, 2006). The important role of AI in empathy was further corroborated by two studies focusing on the effects of expertise in compassion. Lutz *et al.* (2004) found greater insula activation in long-time meditators (Buddhist monks) as compared to untrained (control) meditators while they were engaging in meditation practices for the cultivation of compassion, which the monks had practiced for many years. Similarly, Lazar *et al.* (2005) found significant differences in cortical thickness in prefrontal cortex and right AI in experienced as compared to non-meditators. Figure 17.1 schematically summarizes the areas (marked in red) found to be typically activated in empathy tasks.

Individual Differences in Empathy

Evidence for individual differences in empathic skills has not only been found in studies of populations with a high degree of empathic skills stemming from long-time compassion meditation, but also in the abovementioned empathy studies of adults randomly selected from the normal population. As we all experience in our everyday lives, people are not equally empathic. Scientifically, individual differences in empathic capacity can be assessed using standard empathy questionnaires developed and validated by psychologists, such as the Empathic Concern Scale of the Interpersonal Reactivity Index (IRI; Davis, 1980) and the Balanced Emotional Empathy Scale (BEES; Mehrabian and Epstein, 1972). These scales measure psychological traits which are conceptualized as personality dispositions that are relatively stable over the lifespan of a given person. Analyses of empathic brain responses obtained while subjects were observing other people suffering – be it their loved ones or people the subjects liked (Singer *et al.*, 2004, 2006) – have revealed individual differences in activity in empathy-related pain-sensitive areas (ACC and AI), and that these differences co-vary with interindividual differences in IRI and BEES scores. The higher subjects scored on these questionnaires, the higher their activation in ACC and anterior insula. Interestingly, Jabbi *et al.* (2007) observed similar correlations between IRI subscales and empathic brain responses in the AI for subjects who had observed others tasting pleasant or unpleasant drinks associated with facial expressions of joy or disgust alternatively. Empathic brain responses are not only positively correlated with trait measures of empathy, but also with unpleasantness ratings which subjects give online after each trial of an empathy-inducing condition of a scanning session (Jackson *et al.*, 2005; Lamm *et al.*, 2007; Saarela *et al.*, 2007). Future research will have to clarify how these individual differences in empathic brain responses come about, and whether they are able to explain individual differences in prosocial behavior – two lines of research which have not yet been sufficiently addressed. In summary, findings of shared circuitries underlying one's own sensations and feelings and the observation of similar sensations and feelings in others suggest that we use neural representations reflecting our own emotional responses to understand how it feels for others to be in a similar state. They further suggest that our ability to empathize may have evolved from a system which represents our own internal feeling states and allows us to predict the affective outcomes of an event for ourselves and for other people (e.g., Singer *et al.*, 2004). Thus, deficits in representing one's own feeling states should result in deficits in empathizing with others, a hypothesis that will be discussed in more detail below.

Furthermore, results from fMRI studies on empathy for taste, disgust, and pain suggest that AI cortices play a crucial role in empathy (see also Figure 17.1). Some authors also refer to insular cortex as the interoceptive cortex (e.g., Craig, 2002) because this brain region is involved in processing a variety of information about internal bodily states, including pain, taste, hunger, thirst, and arousal. In the following, a more detailed account of the possible functions of interoceptive cortex and its role for feelings in general and empathy in particular is provided.

The Role of Interoceptive Cortex in Feeling and Empathy

Whereas the beginning of affective and social neuroscience was characterized by a strong focus on exploration of the role of amygdala in emotional processing, that focus has now broadened to include another structure that plays a crucial role in processing feelings: the insular cortex and, in particular, the anterior insular (AI) cortex. It has been suggested that these regions represent a crucial part of the human interoceptive cortex (Craig, 2002) and subserve neural representations of internal bodily and feeling states (Damasio, 1994; Critchley, *et al.*, 2001, 2004).

Interoceptive models of emotions have a long tradition in psychology, and propose that cortical representations of internal bodily signals are at the origin of feeling states. In the late eighteenth century, William James and Carl Lange suggested with the now famous *James-Lange theory* that changes in bodily responses are a necessary condition for emotional experience to arise (James, 1894; Lange, 1885). Thus, we feel our hearts beating when we fall in love or experience fear; we feel our stomachs constricting when we are under stress and have to make a difficult decision; and we feel our face reddening with rage or blushing when we experience an embarrassing situation. They argued that emotions cannot be experienced in the absence of these bodily feelings.

Based on anatomical observations in non-human species, Bud Craig has elaborated on these notions and developed a detailed anatomical model suggesting that an image of the body's internal state is first mapped to the brain by afferents that provide input to thalamic nuclei, sensorimotor cortices, and posterior dorsal insula. In humans, this modality-specific sensory representation of the body's physiological condition in the posterior insula is initially re-represented in the anterior insula on the same side of the brain, and then, by way of a callosal pathway, remapped to the other side of the brain in the right AI. Such a second-order re-representation in right AI is assumed to subserve subjective feelings, and was even proposed as the seat of our awareness of a physical self as a feeling entity (see also Damasio, 1994; Critchley *et al.*, 2001). At the same time, afferents also project by way of the medial dorsal thalamic nucleus to ACC to produce a motivation to engage in relevant behavior. Thus, direct activation of both the insula and the ACC may correspond to simultaneous generation of both a feeling and an affective motivation, with its attendant autonomic effects.

Indeed, imaging studies focusing on the relationship between peripheral measures of arousal and brain activity give robust evidence for the crucial role of rostral ACC and AI cortices in the representation of internal bodily states of arousal as well as the awareness of these states (Critchley *et al.*, 2001, 2003, 2004). The role of AI in interoceptive awareness was specifically highlighted by two studies conducted by Critchley's group. To study the effects of peripheral arousal feedback to the brain, they selected subjects with pure autonomic failure (PAF), which entails an inability to generate autonomic arousal due to specific peripheral denervation of the autonomic system. Using a fear-conditioning paradigm, they compared the brain responses of these subjects to those of normal controls when participants either consciously or unconsciously processed angry faces that had been paired with loud, aversive noise stimuli. The control subjects, in contrast to the PAF subjects, showed an autonomic response when exposed to the conditioned emotional stimuli; namely, enhanced activity in right AI. This suggests a sensitivity of right AI to autonomic feedback, which is absent in individuals with PAF. In addition, emotional awareness of the stimuli was manipulated using backward masking, a procedure in which two stimuli are presented in such rapid sequence that the second stimulus "masks" the first, thus disrupting the observer's opportunity to consciously evaluate the content of the first stimulus. In accordance with the theory suggesting a role for AI in the conscious experience of emotions, the researchers demonstrated, as in previous studies, an enhanced BOLD response in amygdala to unconsciously perceived threat stimuli, but only enhanced activation in AI when the conditioned face was consciously perceived (Critchley *et al.*, 2002). In a subsequent study, Critchley demonstrated that the activity and size of right AI were positively associated with the degree to which participants were aware of their own heartbeat (Critchley *et al.*, 2004). Overall, these and other findings suggest that interoceptive cortex plays an important role for the representation and awareness of feeling states arising from the body.

As the abovementioned results on empathic brain responses suggest, the very same structures (AI and ACC) which play a crucial role in representing our own feeling states also seem to be crucial in processing vicarious feelings. Based on this observation, Singer *et al.* (2004) extended an interoceptive model of emotions to the domain of empathy, and suggested that cortical re-representations in AI of bodily states may have a dual function. First, they allow us to form subjective representations of feelings. These representations allow us not only to understand our feelings when emotional stimuli are present, but also to predict the bodily effects of anticipated emotional stimuli to our bodies. Second, they may serve as the visceral correlate of a prospective empathic simulation of how something may feel for others. This may then help us to understand the emotional significance of a particular stimulus and its likely consequences. In accordance with this view, it is noteworthy that, using fMRI, the anticipation of pain has been found to activate more anterior insular regions, whereas the actual experience of pain also activates more posterior insular regions, which confirms the postulated role of more posterior insular regions in modality-specific, primary representations of pain, and more anterior regions in the secondary representations of the anticipatory negative affect related to pain (Ploghaus *et al.*, 1999). Similarly, in Singer *et al.*'s (2004) empathy study, activity in posterior insular cortices – contralateral to the stimulated hand – was only observed when participants were actually experiencing pain themselves, whereas activity in AI was observed when participants were experiencing pain themselves and when they were vicariously simulating it for others.

A model suggesting that the representation of one's own feeling states is necessary for empathy to arise would make two predictions. First, training the capacity to understand our own feelings would go hand in hand with training the capacity for empathy. Second, deficits in understanding one's own emotions should be associated with empathy deficits. Whereas evidence for the first hypothesis is still lacking, evidence for the second hypothesis is slowly accumulating. The next section of this chapter will focus on the role played by interoceptive cortex in pathologies associated with a lack of empathy and social cognition, such as autism and psychopathy.

Understanding Others in Psychopathology: Psychopathy, Autism, and Alexithymia

The study of patients with specific psychopathological deficits in social cognition is relevant not only for the development of effective treatments for these conditions, but also for a better understanding of the mechanisms underlying social cognition in the healthy population. For example, at the beginning of this chapter, I stated that mentalizing and empathizing represent two different capacities that rely on different circuitries, and both result in a better understanding of other people's minds. Figure 17.1 illustrates the two different circuitries, marked in blue and red respectively. These two capacities usually work together, so normal individuals presumably activate both networks when confronted with tasks in which they have to draw inferences about the mental states of others. It would be difficult to prove a dissociation between these two systems. However, one could elegantly support the claim that there are two different pathways for mentalizing and empathizing by demonstrating a double dissociation in two different patient populations, such as autism and psychopathy.

Generally defined, autistic spectrum disorders (ASD) are pervasive developmental disorders characterized by abnormalities of social interaction, impairments in verbal and non-verbal communication, and a restricted repertoire of interests and activities. People with high functioning autism or Asperger syndrome differ from ASD patients mainly in that they have high intelligence and no impairments in verbal communication. People with ASD and Asperger syndrome have often been found to have difficulties in attributing mental states, such as beliefs, desires, or intentions, to others (reviewed by Frith, 2004). It is less clear, however, whether people with Asperger syndrome also lack an ability to empathize with others. Psychopathy, in contrast, is characterized as a personality disorder mainly marked by a lack of empathy, narcissism, impulsivity, selfishness, the instrumental use of others, and altered emotional sensitivity (for an overview, see Hare, 2006). However, it is not clear whether psychopaths also have deficits in their ability to mentalize. Thus, behavioral studies with psychopathic populations have found selective emotional dysfunction, such as impairment in aversive conditioning, autonomic response to threat, augmentation of the startle reflex to visual threat primes, impaired processing, and altered autonomic response to sad and fearful facial and vocal expressions (Hare, 1982; Levenston *et al.*, 2000; Blair, 2001; Flor *et al.*, 2002). Neurophysiological evidence has been reported of reduced amygdala response during emotional memory and aversive-conditioning tasks, as well as reduced gray matter in the amygdala and hippocampus (Kiehl *et al.*, 2001; Veit *et al.*, 2002; Blair, 2005). More importantly, Sterzer *et al.* (2007) investigated the structure of brains of adolescents with conduct disorder, a psychiatric disorder which often develops into antisocial personality disorder after the age of 18 and is associated with over-aggressive behavior towards people, objects, and animals, and with frequent norm violations. Sterzer and colleagues found reduced activity in amygdala and AI in the adolescents with conduct disorder as compared to normal controls. Moreover, the empathy scores of adolescents with conduct disorder correlated negatively with gray-matter volume in bilateral AI, suggesting the important role of this structure for the correct representation of affective state and the subsequent appropriate empathic response. Birbaumer *et al.* (2005) has also reported hypoactivation in amygdala and AI in psychopathic individuals during a fear-conditioning task.

Findings such as these suggest a difficulty in emotional empathic responding without a clear indication of theory-of-mind impairment. As suggested above, these findings are indications of a possible double dissociation of empathizing and mentalizing skills in patients with autism and those with psychopathy. Patients with autism may lack mentalizing abilities, but may not be deficient in the ability to share others' feelings; in contrast, patients with psychopathy lack empathy, but are unimpaired in their understanding of other people's thoughts and beliefs. However, no fMRI study has ever demonstrated a clear dissociation between mentalizing and empathizing deficits in psychopathy and autism.

Our group has started to investigate a similar dissociation by studying interoceptive awareness and empathy in high-functioning patients with autism or Asperger syndrome (AS) and a condition named alexithymia. Whereas autistic spectrum disorders (ASD) are associated with severe deficits in social cognition and communication, alexithymia is a subclinical phenomenon involving a lack of emotional awareness or, more specifically, difficulty in identifying and describing feelings, and in distinguishing feelings from the bodily sensations of emotional arousal (Nemiah *et al.*, 1976). Alexithymia is thought to be present in about 10% of the general population (Linden *et al.*, 1994; Salminen *et al.*, 1999), and was observed in about 50% of high-functioning patients with autism or AS (Hill *et al.*, 2004). Even though we know that people with AS very often have difficulty attributing mental states to others, and that these deficits are associated with less activation in the mentalizing brain network described above (Frith and Frith, 2006), we know very little about their capacity to experience feelings from their own bodies or to empathize with the feelings of others. Therefore, in a recent fMRI study, we scanned subjects with AS and controls with varying degrees of alexithymia while they performed a task that required them to experience their own feelings (Silani *et al.*, 2008). Specifically, subjects were asked to judge how they felt about emotionally arousing pictures. Results showed that the degree of severity in alexithymia, as measured by two different alexithymia scales, was correlated with less activation in anterior insula. However, a lack of activation in insular cortices during interoceptive awareness of emotions was not specific to the AS diagnosis, but was predicted entirely by the degree of alexithymia. Thus, controls with stronger alexithymic symptoms also showed less activation in interoceptive cortex. These data again indicate that AI plays a role in understanding one's own emotions. Interestingly, individual differences in the degree of alexithymia correlated negatively with individual differences in a questionnaire measuring people's empathic personality disposition, and levels of both alexithymia and empathy were predictive of brain activation in AI during interoception. These findings are perfectly in line with the prediction that deficits in understanding one's own emotions result in empathy deficits, and that both should be correlated with lesser activation in AI. Current investigations in our laboratory using a paradigm for the measurement of empathic brain responses to others' suffering and a classical mentalizing paradigm are extending these findings by scanning alexithymic and non-alexithymic subjects with and without a diagnosis of AS to show empathy deficits on the level of brain responses, as well as a double dissociation between empathy and theory of mind. Such a double dissociation would indicate that a lack of theory of mind, but not empathy, is characteristic of AS, but not alexithymia; and that a lack of interoceptive awareness and empathy is characteristic of alexithymia, but not AS without emotional deficits. This finding would be the first to disprove the implicit general notion of a global empathy deficit in autism, and would also provide clear evidence for the existence of two distinct roots to the understanding of other people's minds: a "cold" cognitive one and a "hot" empathic one.

When do we Care About Others? Modulatory Factors of Empathy

In the previous section of this chapter, examples were given from patient populations with severe social and empathy deficits. However, as described above, there are also substantial individual differences with respect to empathy in the normal healthy population. In addition, the degree to which we have empathic feelings also varies as a function of situational factors. For example, it is usually easier to empathize with someone who has treated one well than with someone who has treated one poorly. In the past 2 years, fMRI studies have embarked on an investigation of the modulatory factors of empathic brain responses. For example, with respect to empathy for pain, a subject's affective link to the other person (Singer *et al.*, 2004, 2006), the subject's appraisal of whether the reason the other person is suffering is justified (Lamm *et al.*, 2007), the frequency of a person's prior exposure to pain-inducing situations (Cheng *et al.*, 2007), and the intensity of the inflicted pain (seeing a needle pricking vs penetrating a muscle, Avenanti *et al.*, 2006) all seem to play a role in the modulation of the magnitude of empathic brain responses.

One example that is particularly relevant for the field of economics is a study conducted by Singer *et al.* (2006) demonstrating evidence for the modulation of empathic brain responses to another person's pain as a function of the perceived fairness of the other person. In this study, male and female volunteers first played a repeated trust game as Mover 1 with two confederates. In the game, Mover 1 can send her endowment of 10 points to Mover 2. Mover 2 can then respond to Mover 1's trust by sending 0 to 10 points back to Mover 1. Each point returned to Mover 1 is tripled. Mover 2 can thus opt for strategies with varying degrees of fairness. In Singer *et al.*'s (2006) experiment, one confederate was instructed to play fairly by reciprocating the subject's trust by returning fair amounts of points; the other was asked to play unfairly by selecting only self-interested choices and responding with no or only minimal returns. After that, an empathy-for-pain paradigm similar to the one reported by Singer *et al.* (2004) was used to measure the subject's empathic brain responses while either the subject or one of the confederates was receiving painful stimulation to her hand. To assess gender differences in empathy and its modulation, both men and women were scanned and paired with either two male or two female players. As in previous empathy studies, empathy-related activation in ACC and AI was observed for both genders when the fair, likeable player was in pain. However, men, but not women, showed an absence of such empathic activity when seeing an unfair player in pain. Instead, men showed increased activation in areas associated with reward (nucleus accumbens), which correlated positively with their desire for revenge as assessed by questionnaires after the scanning session. These results suggest that, at least in men, a desire for revenge won over empathic motivation when they were confronted with someone experiencing pain who they believed deserved to be punished. This finding is in agreement with results from a study conducted by de Quervain and colleagues (2004) showing similar reward-related activation when players were scanned while they were able to deliver punishment points to participants who had defected on them in previous games.

This pattern of results contributes to a microfoundation for theories of social preferences. These theories suggest that people's valuations of other players' payoffs depend on how fairly the other players have played in previous games (Fehr and Gächter, 2000): People tend to place a positive value on others' payoffs if the others have played fairly, but a negative value on others' payoffs if the others have played unfairly. This pattern of preferences implies that people prefer to cooperate with fair opponents and to punish unfair opponents. And indeed, a variety of experiments in behavioral economics were able to demonstrate that people are willing to punish players who had previously behaved unfairly in monetary exchange games, even though this behavior may seem irrational because they knew they would never see the other player again and had to pay to punish the other player. They were therefore working against optimizing their own income by spending fair amounts of money on punishing other people. Economists term this behavior *altruistic punishment*, due to its costly nature (for details, see Chapter 15 of this volume). Neuroscientific findings suggest that punishing people who have violated social norms activates reward circuitries usually engaged in processing primary rewards. This may explain why people are motivated to engage in altruistic punishment even if this behavior is costly: it is rewarding.

Further investigation of the factors that modulate empathic brain responses will be of great relevance for a better understanding of the conditions under which prosocial and other-regarding behavior, on the one hand, and revenge-driven or egoistic behavior, on the other, are more likely to occur.

OPEN QUESTIONS AND IMPLICATIONS FOR FUTURE RESEARCH IN NEUROECONOMICS

Recent efforts in social neuroscience and neuroeconomics have helped shed light on the mechanisms underlying our ability to mentalize and to empathize. Even though these two abilities have distinct features and seem to rely on different neuronal circuitries, both allow humans to represent other people's states – their intentions, beliefs, and thoughts, their emotions and sensations. In the following section, important open questions will be raised and discussed in light of their implications for economic theory and future neuroeconomics studies.

The Automaticity Assumption and its Relationship to Social Preferences

One important feature of the mechanisms outlined is that representing other people's states and goals seems to happen mostly automatically and without awareness. For example, in the earlier empathy studies in the domains of pain, touch, and disgust, subjects were not even told that the goal of the study was to investigate empathic brain responses, but merely instructed passively to watch a scene or movie (Wicker *et al.*, 2003; Keysers *et al.*, 2004; Singer *et al.*, 2004, 2006). Nevertheless, even without explicit instruction, when we perceive others' feelings, the brain networks that represent our own feelings appear to be automatically activated. It seems that we automatically share other people's feelings. This observation is in line with earlier perception–action models of motor behavior and imitation, and with their extension to the domain of empathy (Preston and de Waal, 2002; Gallese, 2003). For example, Preston and de Waal (2002) proposed a neuroscientific model of empathy suggesting that observation or imagination of another person in a particular emotional state *automatically* activates a representation of that state in the observer with its associated autonomic and somatic responses. The term "automatic" in this case refers to a process that does not require conscious and effortful processing, but which can nevertheless be inhibited or controlled.

Such automatic perception–action mechanisms could have important implications for economic theory. Without thinking about or being consciously aware of doing so, we perceive others' feelings and motivational states; this activates the same brain networks that represent our own feelings and motivational states. Empathic resonance mechanisms such as these create a link between our own and other people's needs, desires, motivational states, and emotions. Since other people's emotions affect our own emotional state, and our own emotions are important determinants of our motives and actions, other people's feeling states will partly shape our own motives towards them. This implies that our motives are not only self-interested, but also automatically other-regarding due to resonance mechanisms such as these.

Figure 17.2 exemplifies how empathic resonance mechanisms may influence economic models for the prediction of actions in social exchange settings as conceived in game theoretical frameworks. Accordingly, our actions in such games are determined by our preferences (e.g., desires, beliefs, needs) and our beliefs about the other player's preferences which, in turn, will determine the other player's actions. This chapter has reviewed evidence from social neuroscience about the mechanisms enabling us to draw inferences about other people's mental and motivational states, as well as about their actions and intentions. Empathic resonance mechanisms may establish a link between one's ability to predict others' motives and the nature of one's own motives. That is, other people's emotions may partly shape our own motives concerning them. This link is illustrated as a red arrow in Figure 17.2. To provide an example: when confronted with a person suffering from her parents' having passed away, many people will empathize with the person; this shared suffering may motivate them to try to alleviate the other person's suffering (and thereby feel better themselves), even though they will incur costs doing so and even though this other-regarding behavior may not have been produced were it not for empathic resonance mechanisms. Donations to charity may draw on such empathic mechanisms, when, for example, people are motivated to donate money to charities that operate in third-world countries after seeing a documentary about starving children.

The Link between Empathy, Fairness, "Theory of Mind," and Prosocial Behavior

Of course, the propensity to engage in other-regarding behavior is not determined exclusively by whether a person is equipped with such empathic resonance mechanisms, but by many other factors as well. One such factor is that emotions and motivational states are transient phenomena, so the chronological proximity of the empathy-inducing stimulus may matter.

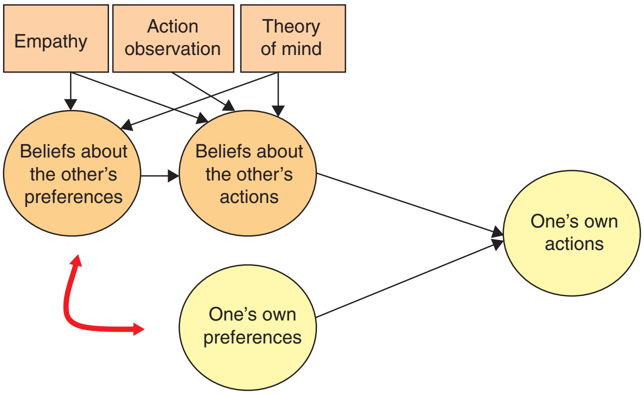


FIGURE 17.2 Economic model for the prediction of actions.

As we saw above, empathic brain responses have also been found to be modulated by factors such as the affective link or perceived fairness of another person. Thus, at least in men, empathic brain responses seem to be absent when people who were previously defecting in economic games are now in pain. These empathic brain responses were even replaced by their opposite: feelings of revenge and *Schadenfreude* – the joy of seeing another person suffering. Based on these and other observations, de Vignemont and Singer (2006) recently questioned the automaticity assumption of empathy, and proposed an alternative model in which empathy is not always merely the automatic consequence of the passive observation of emotional cues in others, but rather subject to an appraisal of the situation and modulation (de Vignemont and Singer, 2006). Accordingly, the magnitude of empathic brain responses is determined by a variety of modulating factors ranging from features associated with the emotional display in the suffering person (e.g., intensity of the emotions), the relationship between the empathizer and the other person (e.g., affective link and interdependence), to the features of the empathizer (e.g., gender, emotional experiences, and emotion regulation ability). A better understanding of the nature of these modulatory factors will also help us to understand the complex interplay of conditions determining prosocial behavior.

One prediction, however, that can easily be made is that people with a greater ability to empathize should display more other-regarding behavior. Even though the question concerning how empathy relates to prosocial behavior is crucial for our understanding of human nature and the development of methods to enhance prosocial behavior, surprisingly, this link has not yet been investigated in the context of neuroeconomics and social neuroscience research. This is partly due to the lack of a clear conceptualization of what we mean when we talk about prosocial behavior and the fact that we still lack adequate tasks for a well-controlled quantitative assessment of individual differences in prosocial behavior in a laboratory – tasks that can just as well be used in experiments performed in economics, neuroscience, and behavioral psychology. Thus, whereas economists have mainly used game theoretical paradigms such as the dictator game, the trust game, and the public good game to operationalize prosocial versus selfish behavior in economic exchange, social psychologists and developmental psychologists have mostly used ecologically valid tasks in which people are observed when confronted with situations in which they could help or refrain from helping. While the former measures allow for a more controlled quantitative assessment of prosocial behavior (e.g., the number of monetary units sent to the other player), they suffer due to a lack of realism and demand effects in that either the task instructions or the task properties clearly indicate to the subjects how they are expected to behave. Thus, in most economic games, subjects are only given a choice between engaging in prosocial/fair or selfish/unfair behavior; this precludes the possibility of observing spontaneous, non-instructed prosocial behavior. More importantly, it is not clear whether such economic tasks assess the same prosocial behavior as the prosocial behavior measured in social psychology tasks that focus on the act of helping people in need. Prosocial behavior as assessed in game theoretical paradigms may well be based on fairness motives rather than empathic motivation. Thus, it may well be that individual differences in empathic brain responses may be a good predictor of helping behavior, but not of fairness-based cooperative behavior.

The distinction between fairness and empathic motivation and their relation to different forms of prosocial behavior has not been investigated to a sufficient extent – neither in the field of economics, nor in social sciences or neurosciences. It should be the focus of future neuroeconomics studies. Even though fairness and empathy may seem to be concepts closely related and important for the engagement in prosocial behavior, having one and not the other motivation may also lead to very different outcomes. Thus, studies in experimental economics have repeatedly provided evidence that fairness preferences lead to cooperative behavior, but also to a desire for revenge and retributivism when violated. People who have been treated unfairly in economic games are willing to punish unfair players (Fehr and Gächter, 2002; Fehr and Fischbacher, 2003), and recent neuroeconomics studies have provided evidence for reward-related activation in the brains of people who either actively (de Quervain *et al*., 2004) or passively (Singer *et al*., 2006) take revenge on these defectors. In contrast, purely empathic motivation may also result in cooperative behavior in similar games but to a lesser extent to revenge-related, that is, other-harming behavior, since this behavior is opposed to an empathic motivation.

Another interesting prediction that can be derived from neuroscientific investigations on empathy and cognitive perspective-taking or theory of mind is that empathic motivation is a better predictor of engagement in other-regarding behavior than cognitive perspective-taking is, because empathy engages motivational and emotional brain circuitries, while theory of mind relies on networks (mPFC, TPF, STS, temporal poles) typically believed to be less relevant for motivation and emotions. As stated above, psychopaths may lack empathy, but not cognitive perspective-taking, explaining why they can engage in antisocial behavior while being very good at manipulating and fooling other people – an ability that requires an understanding of other people's beliefs and intentions.

Even though individual differences in empathy may be better predictors of the engagement in prosocial behavior, both capacities, emotional and cognitive perspective-taking, should help people to better predict other people's actions. An interesting question for future research is to determine the relative importance of our ability to empathize and to mentalize for the prediction of others' motives and actions in different situations, and to determine in which situations one is interfering with or even facilitating the other.

References

Adolphs, R. (1999). Social cognition and the human brain. *Trends Cogn. Sci.* 3, 469–479.

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nat. Rev. Neurosci.* 4, 165–178.

Avenanti, A., Bueti, D., Galati, G., and Aglioti, S. (2005). Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain. *Nat. Neurosci.* 8, 955–960.

Avenanti, A., Paluello, I.M., Bufalari, I., and Aglioti, S.M. (2006). Stimulus-driven modulation of motor-evoked potentials during observation of others' pain. *NeuroImage* 32, 316–324.

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

Baron-Cohen, S., Leslie, A.M., and Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition* 21, 37–46.

Batson, C.D. (1987). Prosocial motivation: is it ever truly altruistic? In: L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, Vol. 20. New York, NY: Academic Press, pp. 65–122.

Batson, C.D. (2008). These things called empathy. In: J. Decety and W. Ickes (eds), *The Social Neuroscience of Empathy*. Cambridge, MA: MIT Press, (in press).

Batson, C.D., Fultz, J., and Schoenrade, P.A. (1987). Distress and empathy: two qualitatively distinct vicarious emotions with different motivational consequences. *J. Personality* 55, 19–40.

Batson, C.D. and Shaw, L.L. (1991). Evidence for altruism: toward a pluralism of prosocial motives. *Psychological Inquiry* 2, 107–122.

Birbaumer, N., Veit, R., Lotze, M. *et al*. (2005). Deficient fear conditioning in psychopathy: A functional magnetic resonance imaging study. *Arch. Gen. Psych.* 62, 799–805.

Blair, R.J. (2001). Neurocognitive models of aggression, the antisocial personality disorders and psychopathy. *J. Neurol. Neurosurg. Psych.* 716, 727–731.

Blair, R.J. (2005). Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness Cogn.* 14, 698–718.

Blakemore, S.J., Winston, J., and Frith, U. (2004). Social cognitive neuroscience: where are we heading? *Trends Cogn. Sci.* 8, 216–222.

Botvinick, M., Jha, A.P., Bylsma, L.M. *et al*. (2005). Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. *NeuroImage* 25, 312–319.

Bufalari, I., Aprile, T., Avenanti, A. *et al*. (2007). Empathy for pain and touch in the human somatosensory cortex. *Cerebral Cortex* 17, 2553–2561.

Call, J. (2007). Past and present challenges in theory of mind research in nonhuman primates. *Prog. Brain Res.* 164, 341–353.

Carr, L., Iacoboni, M., Dubeau, M.-C. *et al*. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proc. Natl Acad. Sci. USA* 100, 5497–5502.

Carruthers, P. and Smith, P. (eds) (1996). *Theories of Theories of Mind*. Cambridge: Cambridge University Press.

Cheng, Y., Lin, C.P., Liu, H.L. *et al*. (2007). Expertise modulates the perception of pain in others. *Curr. Biology* 17, 1708–1713.

Clements, W.A. and Perner, J. (1994). Implicit understanding of belief. *Cogn. Dev.* 9, 377–395.

Craig, A.D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3, 655–666.

Critchley, H.D., Mathias, C.J., and Dolan, R.J. (2001). Neuroanatomical basis for first- and second-order representations of bodily states. *Nat. Neurosci.* 4, 207–212.

Critchley, H.D., Mathias, C.J., and Dolan, R.J. (2002). Fear conditioning in humans: the influence of awareness and autonomic arousal on functional neuroanatomy. *Neuron* 33, 653–663.

Critchley, H.D., Mathias, C.J., Josephs, O. *et al*. (2003). Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. *Brain J. Neurol.* 126, 2139–2152.

Critchley, H.D., Wiens, S., Rotshtein, P. *et al*. (2004). Neural system supporting interoceptive awareness. *Nat. Neurosci.* 7, 189–195.

Damasio, A.R. (1994). Descartes' error and the future of human life. *Scientific Am.* 271, 144.

Dapretto, M., Davies, M.S., Pfeifer, J.H. *et al*. (2006). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nat. Neurosci.* 9, 28–30.

Davis, M.H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Cat. Selected Doc. Psychol.* 10.

Decety, J. and Jackson, P.L. (2004). The functional architecture of human empathy. *Behav. Cogn. Neurosci. Rev.* 3, 71–100.

Decety, J. and Lamm, C. (2006). Human empathy through the lens of social neuroscience. *Scientific World J.* 6, 1146–1163.

Decety, J. and Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *Neuroscientist*, [doi:10.1177/1073858407304654].

Dennett, D.C. (1978). Beliefs about beliefs. *Behav. Brain Sci.* 1, 568–570.

de Quervain, D.J., Fischbacher, U., Treyer, V. *et al*. (2004). The neural basis of altruistic punishment. *Science* 305, 1254–1258.

de Vignemont, F. and Singer, T. (2006). The empathic brain: how, when and why? *Trends Cogn. Sci.* 10, 435–441.

Eisenberg, N. and Fabes, R.A. (1990). Empathy: conceptualization, measurement, and relation to prosocial behavior. *Motiv. Emotion* 14, 131–149.

Eisenberg, N. and Morris, A.S. (2001). The origins and social significance of empathy-related responding. A review of empathy and moral development: Implications for caring and justice by M.L. Hoffman. *Social Justice Res.* 14, 95–120.

Fehr, E. and Fischbacher, U. (2003). The nature of human altruism. *Nature* 425, 785–791.

Fehr, E. and Gächter, S. (2000). Fairness and retaliation: the economics of reciprocity. *J. Econ. Persp.* 14, 159–181.

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature* 415, 137–140.

Ferrari, P.F., Gallese, V., Rizzolatti, G., and Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur. J. Neurosci.* 17, 1703–1714.

Flor, H., Birbaumer, N., Hermann, C. *et al*. (2002). Aversive Pavlovian conditioning in psychopaths: peripheral and central correlates. *Psychophysiology* 39, 505–518.

Fogassi, L., Ferrari, P.F., Gesierich, B. *et al*. (2005). Parietal lobe: from action organization to intention understanding. *Science* 308, 662–667.

Frith, C.D. (2004). Emanuel Miller Lecture: Confusions and controversies about Asperger syndrome. *J. Child Psychol. Psych.* 45, 672–686.

Frith, C.D. and Frith, U. (1999). Interacting minds: a biological basis. *Science* 286, 1692–1695.

Frith, C.D. and Frith, U. (2006). The neural basis of mentalizing. *Neuron* 50, 531–534.

Frith, U. (2001). Mind blindness and the brain in autism. *Neuron* 32, 969–979.

Frith, U. and Frith, C.D. (2003). Development and neurophysiology of mentalizing. *Phil. Trans. R. Soc. Lond. B* 358, 459–473.

Gallagher, H.L. and Frith, C.D. (2003). Functional imaging of “theory of mind”. *Trends Cogn. Sci.* 7, 77–83.

Gallagher, H.L., Jack, A.I., Roepstorff, A., and Frith, C.D. (2002). Imaging the intentional stance in a competitive game. *NeuroImage* 16, 814–821.

Gallese, V. (2003). The manifold nature of interpersonal relations: the quest for a common mechanism. *Phil. Trans. R. Soc. Lond. B* 358, 517–528.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609.

Gopnik, A. and Wellman, H. (1994). The “theory theory”. In: L. Hirschfield and S. Gelman (eds), *Domain Specificity in Culture and Cognition*. New York, NY: Cambridge University Press, pp. 257–293.

Gordon, R. (1986). Folk psychology as simulation. *Mind Language* 1, 158–171, Reprinted in 1995 in M. Davies and T. Stone (eds), Folk Psychology: The Theory of Mind Debate. Oxford: Blackwell Publishers.

Greene, J.D. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends Cogn. Sci.* 11, 322–323, author reply 323–324.

Greene, J.D., Sommerville, R.B., Nystrom, L.E. *et al*. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105–2108.

Greene, J.D., Nystrom, L.E., Engell, A.D. *et al*. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44, 389–400.

Grezes, J. and Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Hum. Brain Mapp.* 12, 1–19.

Gu, X. and Han, S. (2007). Attention and reality constraints on the neural processes of empathy for pain. *NeuroImage* 36, 256–267.

Hare, R.D. (1982). Psychopathy and physiological activity during anticipation of an adversive stimulus in a distraction paradigm. *Psychophysiology* 19, 266–271.

Hare, R.D. (2006). Psychopathy: a clinical and forensic overview. *Psych. Clin. North Am.* 29, 709–724.

Harris, P. (1989). *Children and Emotion*. Oxford: Blackwell Publishers.

Harrison, N.A., Singer, T., Rotshtein, P. *et al*. (2006). Pupillary contagion: central mechanisms engaged in sadness processing. *Social Cogn. Affect. Neurosci.* 1, 5–17.

Hart, A.J., Whalen, P.J., Shin, L.M. *et al*. (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *NeuroReport* 11, 2351–2355.

Heal, J. (1986). Replication and functionalism. In: J. Butterfield (ed.), *Language, Mind, and Logic*. Cambridge: Cambridge University Press, Cambridge, Reprinted in 1995 in M. Davies and T. Stone (eds), Folk Psychology: The Theory of Mind Debate. Oxford: Blackwell Publishers.

Hill, E., Berthoz, S., and Frith, U. (2004). Cognitive processing of own emotions in individuals with autistic spectrum disorders and in their relatives. *J. Autism Dev. Disorders* 34, 229–235.

Hoffman, M.L. (2000). *Empathy and Moral Development: Implications for Caring and Justice*. Cambridge: Cambridge University Press.

Iacoboni, M. and Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nat. Rev. Neurosci.* 7, 942–951.

Jabbi, M., Swart, M., and Keysers, C. (2007). Empathy for positive and negative emotions in the gustatory cortex. *NeuroImage* 34, 1744–1753.

Jackson, P.L., Meltzoff, A.N., and Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *NeuroImage* 24, 771–779.

Jackson, P.L., Brunet, E., Meltzoff, A.N., and Decety, J. (2006). Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia* 44, 752–761.

Jacob, P. and Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends Cogn. Sci.* 9, 21–25.

James, W. (1894). Physical basis of emotion. *Psychological Rev.* 101, 2005–2010.

Keysers, C. and Gazzola, V. (2006). Towards a unifying neural theory of social cognition. *Prog. Brain Res.* 156, 379–401.

Keysers, C. and Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends Cogn. Sci.* 11, 194–196.

Keysers, C., Wicker, B., Gazzola, V. *et al*. (2004). A touching sight: SII/PV activation during the observation and experience of touch. *Neuron* 42, 335–346.

Kiehl, K.A., Smith, A.M., and Hare, R.D. *et al*. (2001). Limbic abnormalities in affective processing by criminal psychopaths as revealed by functional magnetic resonance imaging. *Biol. Psychiatry* 50, 677–684.

King-Casas, B., Tomlin, D., Anen, C. *et al*. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83.

Knoch, D., Pascual-Leone, A., Meyer, K. *et al*. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314, 829–832.

Lamm, C., Batson, C.D., and Decety, J. (2007). The neural substrate of human empathy: effects of perspective taking and cognitive appraisal. *J. Cogn. Neurosci.* 19, 42–58.

Lange, C. (1885). *The Emotions*. Reprinted in 1967 in C. Lange and W. James (Eds), *The Emotions*. New York, NY: Harner Publishing Co.

Lazar, S.W., Kerr, C.E., Wasserman, R.H. *et al*. (2005). Meditation experience is associated with increased cortical thickness. *NeuroReport* 16, 1893–1897.

Levenston, G.K., Patrick, C.J., Bradley, M.M., and Lang, P. (2000). The psychopath as observer: emotion and attention in picture processing. *J. Abnormal Psychol.* 109, 373–386.

Linden, W., Wen, F., and Paulhaus, D.L. (1994). Measuring alexithymia: reliability, validity, and prevalence. In: J. Butcher, C. Spielberger, and N.J. Hillsdale (eds), *Advances in Personality Assessment*. Mahwah, MJ: Lawrence Erlbaum, pp. 125–143.

Lutz, A., Brefczynski-Lewis, J.A., and Davidson, R.J. (2004). Loving-kindness and compassion meditation results in unique patterns of fMRI activation and enhances the reactivity of the insula/cingulate neural circuitry to negative stimuli in meditators. Slide presentation presented at the 34th Annual Meeting of the Society for Neuroscience, San Diego, CA.

McCabe, K. and Singer, T. (2008). Brain signatures of social decision making. In: C. Engel and W. Singer (eds), *Better Than Conscious: Exploiting the Capacity of Humans to Reach Decisions by Both Serial and Parallel Processing of Information*. Cambridge, MA: MIT Press, pp. 103–122.

McCabe, K., Houser, D., Ryan, L. *et al*. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl Acad. Sci. USA* 98, 11832–11835.

Mehrabian, A. and Epstein, N. (1972). A measure of emotional empathy. *J. Personality,* 40, 525–543.

Mitchell, J.P., Heatherton, T.F., and Macrae, C.N. (2002). Distinct neural systems subserve person and object knowledge. *Proc. Natl Acad. Sci. USA* 99, 15238–15243.

Mitchell, J.P., Banaji, M.R., and Macrae, C.N. (2005a). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *NeuroImage* 28, 757–762.

Mitchell, J.P., Banaji, M.R., and Macrae, C.N. (2005b). The link between social cognition and self-referential thought in the medial prefrontal cortex. *J. Cogn. Neurosci.* 17, 1306–1315.

Mitchell, J.P., Macrae, C.N., and Banaji, M.R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron* 50, 655–663, comment 531–534.

Moll, J. and de Oliveira–Souza, R. (2007). Moral judgments, emotions and the utilitarian brain. *Trends Cogn. Sci.* 11, 319–321.

Moll, J., de Oliveira-Souza, R., Bramati, I.E., and Grafman, J. (2002a). Functional networks in emotional moral and nonmoral social judgments. *NeuroImage* 16, 696–703.

Moll, J., de Oliveira-Souza, R., Eslinger, P.J. *et al*. (2002b). The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *J. Neurosci.* 22, 2730–2736.

Moll, J., Krueger, F., Zahn, R. *et al*. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc. Natl Acad. Sci. USA* 103, 15623–15628.

Montague, P.R., Berns, G.S., Cohen, J.D. *et al*. (2002). Hyperscanning: simultaneous fMRI during linked social interactions. *NeuroImage* 16, 1159–1164.

Moriguchi, Y., Decety, J., Ohnishi, T. *et al*. (2007). Empathy and judging other's pain: an fMRI study of alexithymia. *Cerebral Cortex* 17, 2223–2234.

Morris, J.S., Frith, C.D., Perrett, D.I. *et al*. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature* 383, 812–815.

Morrison, I. and Downing, P.E. (2007). Organization of felt and seen pain responses in anterior cingulate cortex. *NeuroImage* 37, 642–651.

Morrison, I., Lloyd, D., di Pellegrino, G., and Roberts, N. (2004). Vicarious responses to pain in anterior cingulate cortex: is empathy a multisensory issue? *Cogn. Affect. Behav. Neurosci.* 4, 270–278.

Morrison, I., Peelen, M.V., and Downing, P.E. (2007). The sight of others' pain modulates motor processing in human cingulate cortex. *Cerebral Cortex* 17, 2214–2222.

Nemiah, J.C., Freyberg, H., and Sifneos, P.E. (1976). Alexithymia: a view of the psychosomatic process. In: O.W. Hill (ed.), *Modern Trends in Psychosomatic Medicine*, Vol. 3. London: Butterworths, pp. 430–439.

Ochsner, K.N. and Lieberman, M.D. (2001). The emergence of social cognitive neuroscience. *Am. Psychologist* 56, 717–734.

O'Doherty, J., Winston, J., Critchley, H. *et al*. (2003). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia* 41, 147–155.

Perner, J.L., Leekam, S.R., and Wimmer, H. (1987). 2-year-olds' difficulty with false belief: the case for a conceptual deficit. *Br. J. Dev. Psychol.* 5, 125–137.

Ploghaus, A., Tracey, I., Gati, J.S. *et al*. (1999). Dissociating pain from its anticipation in the human brain. *Science* 284, 1979–1981.

Povinelli, D.J. and Bering, J.M. (2002). The mentality of apes revisited. *Curr. Dir. Psychol. Sci.* 11, 115–119.

Premack, D. and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526.

Preston, S.D. and de Waal, F.B. (2002). Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* 25, 1–20.

Prinz, W. (1990). A common-coding approach to perception and action. In: O. Neumann and W. Prinz (eds), *Relationships Between Perception and Action: Current Approaches*. Berlin: Springer, pp. 167–201.

Rilling, J.K., Gutman, D., Zeh, T. *et al*. (2002). A neural basis for social cooperation. *Neuron* 35, 395–405.

Rilling, J.K., Sanfey, A.G., Aronson, J.A. *et al*. (2004). The neural correlates of theory of mind within interpersonal interactions. *NeuroImage* 22, 1694–1703.

Rilling, J.K., Glenn, A.L., Jairam, M.R. *et al*. (2007). Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biol. Psychiatry* 61, 1260–1271.

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* 3, 131–141.

Saarela, M.V., Hlushchuk, Y., Williams, A.C. *et al*. (2007). The compassionate brain: humans detect intensity of pain from another's face. *Cerebral Cortex* 17, 230–237.

Salminen, J.K., Saarijarvi, S., Aarela, E. *et al*. (1999). Prevalence of alexithymia and its association with sociodemographic variables in the general population of Finland. *J. Psychosomatic Res.* 46, 75–82.

Sanfey, A.G., Rilling, J.K., Aronson, J.A. *et al*. (2003). The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1758.

Saxe, R. (2005). Against simulation: the argument from error. *Trends Cogn. Sci.* 9, 174–179.

Saxe, R. and Powell, L.J. (2006). It's the thought that counts: apecific brain regions for one component of theory of mind. *Psychological Sci.* 17, 692–699.

Saxe, R. and Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia* 43, 1391–1399.

Silani, G., Bird, G., Brindley, R. *et al*. (2008). Levels of emotional awareness and autism: an fMRI study. *Social Neurosci.*, DOI 10.1080/17470910701577020.

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neurosci. Biobehav. Rev.* 30, 855–863.

Singer, T. and Fehr, E. (2005). The neuroeconomics of mind reading and empathy. *Am. Econ. Rev.* 95, 340–345.

Singer, T., Kiebel, S.J., Winston, J.S., Dolan, R.J., and Firth, C.D. (2004). Brain responses to the acquired moral status of faces. *Neuron* 41, 653–662.

Singer, T., Seymour, B., O'Doherty, J. *et al*. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science* 303, 1157–1162.

Singer, T., Seymour, B., O'Doherty, J.P. *et al*. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439, 466–469.

Smith, A. (2004). *The Theory of Moral Sentiments*. Kila, MT: Kessinger Publishing, (originally published in 1759).

Sterzer, P., Stadler, C., Poustka, F., and Kleinschmidt, A. (2007). A structural neural deficit in adolescents with conduct disorder and its association with lack of empathy. *NeuroImage* 37, 335–342.

Veit, R., Flor, H., Erb, M. *et al*. (2002). Brain circuits involved in emotional learning in antisocial behavior and social phobia in humans. *Neurosci. Letts* 3283, 233–236.

Wellman, H. (1990). *The Child's Theory of Mind*. Cambridge, MA: MIT Press.

Wicker, B., Keysers, C., Plailly, J. *et al*. (2003). Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron* 40, 655–664.

Wimmer, H. and Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–128.

Winston, J.S., Strange, B.A., O'Doherty, J., and Dolan, R.J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nat. Neurosci.* 5, 277–283.

Wispe, L. (1986). The distinction between sympathy and empathy: to call forth a concept, a word is needed. *J. Pers. Social Psychol.* 50, 314–321.

At. Quate ver adit wisl utpat, sed tie molut exerit irit accum dolut adiamconsed dolortion velit dipsum ad moloreet, consed magna commodolore commod dolorer sent wissed mod eugiamet,

CHAPTER

18

Social Preferences in Primates

Joan B. Silk

OUTLINE

> When I do good, I feel good; when I do bad,I feel bad, and that is my religion.
>
> Abraham Lincoln

INTRODUCTION

Humans, like all of the other organisms on earth, are the product of evolution by natural selection. Natural selection generally favors traits that enable individuals to survive and reproduce successfully. Success in what Charles Darwin (1859) called "the struggle for existence" often comes at the expense of others, leaving little scope for the kinds of prosocial sentiments that Lincoln expressed. Darwin himself was troubled by the fact that his theory could not explain the altruistic existence of sterile workers in social insect colonies, which spend their lives caring for the offspring of the queen, but never reproduce themselves. In *The Origin of Species*, he wrote that the altruism of social insects presented “one special difficulty, which at first appeared to me insuperable, and actually fatal to my theory.” This conundrum has engrossed evolutionary biologists for the last 40 years, and has generated a large theoretical and empirical literature on the evolution of altruism. This work shows that altruism is widespread in nature, but is typically limited to kin and reciprocating partners (Crozier and Pamilo, 1996; Dugatkin, 1997; Kappeler and van Schaik, 2006). This literature has also illuminated the gap between humans and other animals. Humans rely on cooperation to a far greater extent than most other animals do, and are able to orchestrate cooperation in substantially larger groups. Moreover, humans are the only animals that regularly provide aid to strangers and impose costly punishment on wrongdoers in anonymous, one-shot interactions (Fehr and Fischbacher, 2003; Boyd and Richerson, 2005; Henrich *et al*. 2006). Cooperation in humans is sustained by a willingness to impose costly punishment on those who shirk social obligations. In humans, altruism seems to be motivated at least in part by social preferences based on empathy, concern for the welfare of others, and a preference for equity (Batson, 1991; Fehr and Fischbacher, 2003; see also Chapter 19 of this volume).

What is the evolutionary source of human prosocial preferences? If our prosocial preferences are based on empathy, which relies on the ability to perceive the thoughts and feelings of others, then comparative studies of the cooperation, cognition, and capacity for empathy in humans and and other primates may provide clues about the origins of prosocial preferences. On the other hand, our prosocial preferences might reflect the evolutionary consequences of the economic importance of cooperation in human societies. If that is the case, then it might be profitable to examine the nature of social preferences in other species in which cooperation plays an important role.

Here, I review what we know about the evolutionary foundation and deployment, of altruism in non-human primate species. I begin with a brief primer on the evolution of altruism, and briefly describe the pattern and scope of altruism among primates in the wild, including the deployment of both beneficent behavior and punishment. Then, I examine what is known about the cognitive capacities that underlie empathy, and evidence for empathy and sympathy in non-human primates. Finally, I review what we know about the nature of social preferences that motivate altruistic behavior in non-human primates, focusing on recent experimental studies that probe the nature of prosocial preferences in chimpanzees.

THE ADAPTIVE CHALLENGE OF ALTRUISM

Biologists define altruism as any behavior that is costly to the actor and beneficial to the recipient. By performing an altruistic behavior, actors incur costs that reduce their own chance of reproducing successfully (fitness), and provide benefits that increase the recipient’s fitness. If altruists provide benefits to others indiscriminately, then the benefits will not increase the relative fitness of altruists. However, altruists always bear the costs. The average fitness of a genetic variant (allele) that increases the likelihood of performing the altruistic behavior will therefore be lower than the average fitness of the non-altruistic allele. In order for altruism to evolve, there must be some process that allows altruists to direct benefits selectively with other altruists. In nature, two types of processes can produce this outcome: nepotism and contingent reciprocity.

Kin Selection

Selection can favor altruism toward close relatives because kinship provides a reliable cue of genetic similarity. W.D. Hamilton realized that individuals that are descended from the same ancestors have some probability of inheriting copies of the same genes. In particular, individuals who carry genes that are associated with altruistic behavior are more likely to have relatives who carry copies of the same genes than individuals drawn at random from the population. If individuals behave altruistically to their relatives, they have some chance of conferring benefits on individuals who also carry copies of the the genes that lead to altruistic behavior. This is the underlying foundation for the theory of kin selection (Hamilton, 1964). What has come to be known as Hamilton’s rule predicts that altruism will be favored when $br > c$. The quantities b and c represent the benefits and costs associated with the altruistic act. The quantity r measures how much the possession of a particular gene in one individual predicts the presence of the same gene in a second individual.

If there is limited movement in and out of groups, levels of genetic relatedness will build up over time. When this is the case, group membership can provide a cue for assortative interaction because genes that generate altruistic behavior are disproportionately likely to be shared by other group members. This holds even for individuals who do not share recent ancestry.

Multi-level selection models (Wilson and Sober, 1994) provide an alternative, but equivalent, description of the same process (Dugatkin and Reeve, 1994; Reeve and

Keller, 1997). In the Hamilton's rule approach, fitness effects are allocated to the bodies in which the genes causing the effects are expressed. In the multi-level selection approach, fitness effects are partitioned into within-group and between-group components. The two approaches are mathematically equivalent, but their heuristic value may vary in different circumstances.

Contingent Reciprocity

The basic logic underlying contingent reciprocity, or reciprocal altruism, is the same as the logic underlying kin selection, but now previous behavior provides a cue about whether others carry alleles that lead to altruistic behaviors. When individuals interact more than once, contingent altruistic strategies (such as tit-for-tat) can arise. (For example, in the one-shot version of the prisoner's dilemma game, two suspects of a crime are apprehended and interrogated separately. Each suspect is offered a reduced sentence if he confesses, and implicates his partner. Not knowing what the other suspect will do, each suspect has a strong incentive to confess. Even though both would be better off if they remained silent, neither can afford to take the chance that the other will confess and implicate him. However, the dynamics of the prisoner's dilemma are altered when two individuals face the same situation repeatedly. Then, contingent strategies may be favored. See McElreath and Boyd (2007) for a more complete discussion of the iterated prisoner's dilemma game.)

In the first interaction, an individual who carries the gene that leads to altruistic behavior provides help, but continues to help only if his partner reciprocates. Thus, after the first interaction, contingent altruists direct their costly help only toward other altruists. These kinds of contingent strategies can be sustained as long as $(1-1/t)b > c$, where b is the benefit derived from the other's helpful act, c is the cost of the helpful act, and t is the expected number of interactions between the two. Note that the inequality cannot be satisfied when $t = 1$, so repeat business is required for contingent reciprocity to be favored. This repeated process is the foundation of the theory of reciprocal altruism, which was first introduced by Robert Trivers (1971) and later formalized by Robert Axelrod and William Hamilton (1981).

THE DEPLOYMENT OF ALTRUISM IN PRIMATE GROUPS

There is an extensive literature on the form and distribution of altruism in non-human primate groups (see edited volumes by Chapais and Berman, 2004; Kappeler and van Schaik, 2006). The most common form of altruistic behavior in primate groups is social grooming, which has important hygienic and social functions. Other forms of altruistic behaviors that occur in various primate species include coalitionary support, in which one individual intervenes on behalf of another in an ongoing agonistic interaction; alarm-calling, in which one individual signals to others that a predator is nearby; alloparental care, in which group members help to carry, protect, and care for dependent offspring; and food-sharing, which ranges from active donations of food items to passive tolerance of others feeding in close proximity.

In-group Biases

Virtually all of the New World monkeys (which range through Central and South America) and Old World monkeys and apes (which range through Africa and Asia) live in social groups (Figure 18.1). The size and structure of social groups vary considerably across species, to include pair-bonded nuclear family units; polygynous groups composed of one adult male, multiple adult females, and immatures; and larger groups composed of multiple adult males, multiple adult females, and immatures. In most cases groups move as cohesive units, but some species,

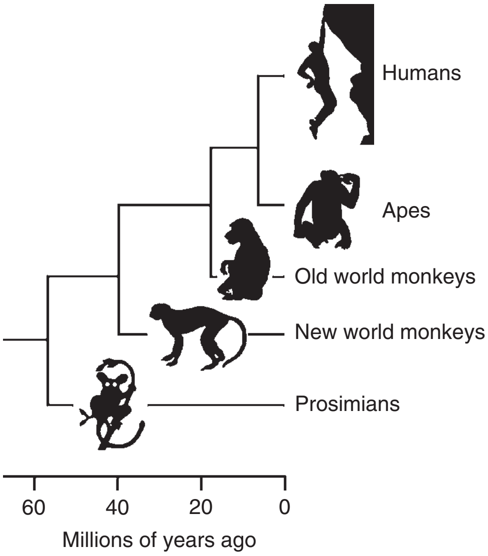


FIGURE 18.1 The primate order is composed of prosimians, New World monkeys, Old World monkeys, and apes. Brain size is generally correlated with body size, but some species have larger brains in relation to their body size than others. In general, apes have relatively larger brains than other primates do.

including chimpanzees, regularly split up into smaller parties. A few primate species form multi-level societies, in which primary social units are organized into larger aggregations that coordinate their ranging behavior and share sleeping sites (Stammbach 1987). (Orang-utans, however, are largely solitary; females associate with their dependent offspring, and juveniles and subadults may form temporary aggregations, but there are no stable associations among adult orang-utans. Mandrills live in groups that may number hundreds of animals, but it is not known whether all group members are recognized as individuals. Similarly, hamadryas and gelada baboons live in multi-level societies, and it is not known whether individual recognition extends to all members of the larger social units.)

Friendly social interactions are restricted to familiar group members. There are very limited exceptions to this rule. For example, females sometimes mate with males from outside their groups (Goodall, 1986; Palombit, 1994; Cords, 2002). Otherwise, responses to strangers and members of neighboring groups range from passive avoidance to active hostility. Some species actively defend their territories from intruders, while others form home ranges and compete with members of neighboring groups for access to resources in areas of range overlap. There is no evidence of cooperation among members of neighboring groups in the wild.

Nepotistic Biases

Dispersal patterns and mating systems influence opportunities for kin-selected altruism. In many primate species, females remain in their natal groups throughout their lives, while males disperse to avoid inbreeding (Pusey, 2004). In most of these species, females form strong and well-differentiated social bonds, and display pronounced maternal kin biases in behavior (Kapsalis, 2004; Silk, 2005). For example, female baboons are significantly more likely to intervene on behalf of their mothers and daughters than for more distantly related females or non-relatives (Silk *et al.*, 2004). They also form longer lasting and more equitable relationships with close maternal kin than with others (Silk *et al.*, 2006a, 2006b). In some species, kin biases extend to paternal kin, although the mechanisms underlying paternal kin recognition are unknown (Widdig, 2007). In chimpanzees, males remain in their natal groups while females disperse, and social bonds among males are well-developed (Muller and Mitani, 2005; Duffy *et al.*, 2007). Males and females show preferences for close maternal kin (Williams *et al* 2002; Langergraber *et al.*, 2007), but maternal kin biases are not as pronounced among male chimpanzees as they are among female monkeys, and paternal kin biases among males are not evident (Langergraber *et al.*, 2007).

The most extreme examples of altruism are found in cooperatively breeding species within the sub-family *Callitrichinae*, the family of New World monkeys that includes marmets and tamarins. In these groups there may be multiple adults of each sex, but breeding is limited to the dominant male and female (Bales *et al.*, 2000). (Dominant females actively enforce their reproductive monopoly. In several cases dominant females have killed infants produced by subordinate females, and subordinate females are only able to rear litters if they gave birth when the dominant female does not have dependent infants; see Digby, 1995). Mature offspring delay dispersal, and normally do not breed in their natal groups. All group members help to care for infants, regardless of their genetic relationship to them.

Kin selection has probably played an important role in the evolution of cooperative breeding in callitrichids. Groups are mainly composed of closely related family members, and helping behavior contributes to their inclusive fitness. Moreover, callitrichids typically produce twins that share a common placenta and chorion (the membrane the surrounds the growing embryo in the uterus). Stem cells are passed from one twin to the other, a process that is called genetic chimerism (Haig, 1999). It has recently been learned that chimerism extends to all bodily tissues, including the gametes (Ross *et al.*, 2007). This means that individuals sometimes pass along their siblings' genes, not their own. Chimerism effectively raises the degree of relatedness within sibling pairs, and may increase the inclusive fitness benefits derived from helping (Haig, 1999).

Altruism Toward Reciprocating Partners

There is considerably less consensus about the role of contingent reciprocity than about the role of nepotism in primate groups (for reviews, see Hammerstein, 2003; Noë, 2005; Silk 2007a). This issue is contentious because it is difficult to demonstrate contingency in natural sequences of behavior.

A series of naturalistic experiments suggest that contingency does influence the pattern of exchanges within dyads. Seyarth and Cheney (1984) showed that wild vervet monkeys were more attentive to the tape-recorded distress calls of unrelated group members if they had been groomed recently by the caller than if they had not been groomed recently by the same monkey. In contrast, grooming among closely related monkeys did not influence the likelihood of responding to distress calls. Similarly, Hemelrijk (1994) showed that long-tailed macaques were more likely to intervene on behalf of monkeys who had recently groomed them than for monkeys who had not groomed them. In both these experiments, researchers used a within-subject design to assess the effects of recent grooming. This allowed them to exclude the possibility that close associates are simply more likely to groom one another and support one another. There have also been a number of more formal experimental studies of contingent cooperation in several species of monkeys (reviewed in Silk, 2007a). These studies generally suggest that altruism by one individual enhances the likelihood of cooperation by the other, but the behavioral strategies and preferences underlying these contingencies are not well established.

PRIMATE POLICING AND PUNISHMENT

Punitive action against potential rivals and competitors is common in nature (Clutton-Brock and Parker, 1995). Thus, a female monkey may attack another female who encroaches on her food patch, or a male may threaten a rival who comes too close to a female that he is mate-guarding. In contrast, humans often punish individuals who violate social norms or fail to cooperate, even when they are not harmed directly themselves. In colonies of social insects, workers routinely destroy eggs laid by rogue workers, preserving the stability and productivity of the entire colony (Ratnieks, 1988; Ratnieks and Wenseleer, 2005). This type of "policing" in humans and social insects is altruistic, because the individual who imposes the punishment incurs costs, while the benefits are widely shared by other group members.

The English language does not provide an easy way to distinguish between punitive action that benefits the individual and punitive action that benefits the group as a whole. For clarity, I use the terms *retaliation* for the former and *altruistic punishment* for the latter. This distinction is important to keep in mind, because the evolutionary forces that underlie them are quite different. Individual selection will favor retaliation because it generates direct benefits for actors (Clutton-Brock and Parker, 1995). Kin selection (multi-level selection processes) may favor altruistic punishment in the highly related colonies of social insects (Ratnieks, 1988; Ratnieks and Wenseleer, 2005), while the combination of cultural group selection and indirect reciprocity may lead to the evolution of altruistic punishment in human societies (Boyd *et al*., 2003; Panchanathan and Boyd, 2004; Gintis *et al*., 2007).

Retaliation is widespread in primate groups. The most recent evidence for retaliation comes from experimental work on chimpanzees (Jensen *et al*., 2007a, 2007b). In the first study, one chimpanzee was given the opportunity to respond to the loss of valued food items by pulling a rope which caused a sliding platform to collapse and the food to fall out of reach. In one condition, a human experimenter moved the platform away from the actor and slid it to within reach of another chimpanzee (Jensen, 2007a). In another condition, the experimenter did the same thing, but there was no other chimpanzee present to receive the food. In the third condition, a chimpanzee in the opposite cage was able to pull the platform away from the actor and gain access to the food. The chimpanzees were more likely to collapse the table when they lost food than when they were left alone to eat in peace. However, the chimpanzees were most likely to become aroused and dump the food when they were victimized by other chimpanzees.

In the second experiment, Jensen and colleagues (2007b) conducted a reduced form of the ultimatum game in which a one chimpanzee (Player 1) was able to choose between two different pre-set distributions of rewards. One option always provided eight pieces for Player 1 and two pieces (8/2) for another chimpanzee (Player 2), while the other option provided a distribution of 5/5, 8/2, or 10/0. To make a choice, Player 1 pulled a rod that was attached to a tray that held the rewards for each animal. To accept Player 1's choice, Player 2 pulled another rod which brought the food rewards to within reach of both individuals, and then Player 1 and Player 2 could claim their respective rewards. If Player 2 did not pull the rod, neither individual got any food. Here, Player 2 had the opportunity to retaliate against Player 1 if unsatisfied with the offer. Although individuals taking the role of Player 1 preferentially chose offers that benefited themselves (8/2 over 5/5), individuals who took the role of Player 2 rarely rejected any non-zero offers, and showed little evidence of arousal in any of the test trials.

There is very limited evidence for altruistic punishment in non-human primates. Two anecdotes suggest that chimpanzees might punish individuals that violate social norms. In the Mahale Mountains of Tanzania, a young adult male was brutally attacked by eight members of his own group (Nishida *et al*., 1995). The authors speculated that this young male may have been victimized because he did not conform to social rules – he did not defer to higher-ranking males, and launched unprovoked attacks on adult females.

More systematic evidence comes from two field experiments conducted on monkeys. Monkeys sometimes give distinctive calls when they find desirable foods (Hauser and Marler, 1993a). Capitalizing on this observation, Hauser and Marler (1993b) surreptitiously provisioned rhesus monkeys (*Macaca mulatta*) with desirable food. In some cases the monkeys who discovered these caches called, and in other cases they remained silent. Female macaques were more likely to be attacked by other group members if they remained silent than if they gave food calls after finding these items. The authors hypothesized that food calls function to announce possession of a food item, and they interpreted the harassment of females who remain silent as a form of punishment for attempting to conceal the location of food items. This would constitute a form of altruistic punishment because the screams of the victim alerted other group members to the site of the food, and many animals had an opportunity to profit from the aggressor's actions.

Subsequent work on food-calling in white-faced capuchins (*Cebus capucinus*; Gros-Louis, 2004) provides an alternate interpretation for aggression in this context. Like rhesus monkeys, capuchins who called after finding food were less likely to be approached by others in the vicinity than were monkeys that remained silent. In addition, individuals who gave food calls when they were approached by higher-ranking animals were less likely to receive aggression than monkeys who did not call. Gros-Louis (2004) suggests that food calls may function to establish the ownership of resources and signal the owners' willingness to defend them, thus deterring potential competitors from trying to take them. This would explain why monkeys are more likely to call when approached by high-ranking monkeys, who might challenge them for possession, than by lower-ranking monkeys, who are unlikely to do so.

In some species of primates, individuals mediate disputes among other group members (Flack *et al.*, 2005). This form of intervention differs from coalitionary aggression, because the actor does not take sides. Impartial mediation may bring disputes to a speedy end, thus reducing the costs incurred by the participants and avoiding the possibility of conflicts escalating. Flack and her colleagues consider mediation a form of policing, and predict that it will only be deployed in species with pronounced asymmetries in power. In these cases, powerful individuals can intervene effectively at minimal cost to themselves. If the costs are negligible, then this form of policing may not represent a form of "altruistic punishment."

COGNITIVE BASIS OF SOCIAL PREFERENCES

In humans, empathy enhances prosocial motivations (Batson, 1991). The capacity for empathy relies on the ability to comprehend the feelings, motives, and thoughts of others, and to appreciate the distinction between one's own thoughts and the thoughts of others (Preston and de Waal, 2002). This, in turn, requires a well-developed theory of mind, a multidimensional construct that includes the ability to attribute perception, attention, desires, goals, intentions, knowledge, and beliefs to others (Call, 2007).

Monkeys and apes have considerable knowledge about social information. For example, they are able to recognize dozens of individuals, identify kin, compute the values of resources and services, keep track of past interactions with group members, make transitive inferences, discriminate between cooperators and defectors, and assess the qualities of prospective rivals, mates, and allies (Cheney and Seyfarth, 2007; Tomasello and Call, 1997). Primates also know something about the nature of dominance, kinship, and affiliative relationships between other group members (Cheney and Seyfarth, 2007).

Monkeys and apes can succeed in tasks that require them to attribute perceptual knowledge to others. For example, chimpanzees use visual gestures selectively when others can see them, and pay particular attention to the orientation of the face (Call, 2007). In experiments in which rhesus monkeys attempt to "steal" food from humans, the monkeys attend to the visual and auditory perceptions of the experimenter (Flombaum and Santos, 2005; Santos *et al.*, 2006).

Experimental studies also provide evidence that chimpanzees have some understanding of others' knowledge and intentions, and can use this information in strategic ways. Hare and colleagues (2000, 2001) created an experimental protocol in which a subordinate chimpanzee was paired with a more dominant group member. The experiment relies on the fact that subordinate individuals are normally reluctant to challenge dominant individuals over access to food rewards. In a central enclosure, food rewards were hidden behind barriers, so that both rewards were visible to the subordinate but only one was visible to the the dominant chimp. The subordinate saw the food being placed in the enclosure, but the dominant did not (Figure 18.2). After the foods were hidden, the chimps were given access to the central enclosure, and the researchers monitored each chimp's movements. They predicted, of course, that the dominant would head directly for the food reward that was visible to it.

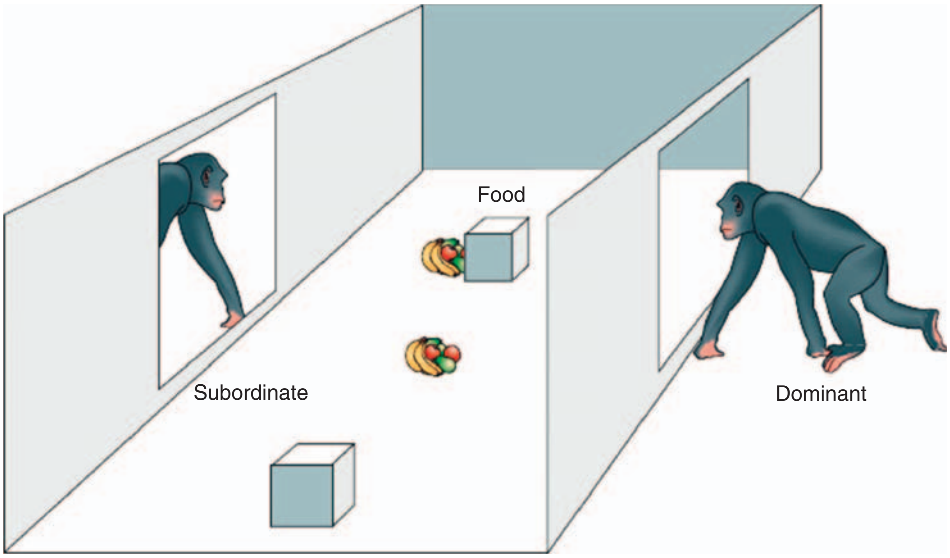


FIGURE 18.2 In this experiment, food rewards were placed in the central area and on one side of an opaque barrier, making both visible to the subordinate (on the left) but only one visible to the dominant. The dominant chimp's knowledge and beliefs about the location of food rewards were varied systematically.Reprinted from Ghazanfar *et al.* (2004), with kind permission of *Nature Review Neuroscience*.

If the subordinate chimp knew what the dominant saw, then it was expected to take advantage of this knowledge and head for the reward that was hidden from the dominant. This is just what the chimpanzees did. So, chimpanzees evidently knew what others knew about the location of food items (see Karin-D'Arcy and Povinelli, 2002, for an alternative interpretation). Capuchin monkeys do not succeed in the same task (Hare *et al.*, 2003).

Comparative studies of apes and children suggest that there are substantial differences in social cognition. While chimpanzees, bonobos, and orangutans do as well as two-and-a-half-year-old children on tasks that require physical cognition (e.g., tracking a reward after it has been moved, using a tool to retrieve a reward that is out of reach), human children are much more successful than apes in tasks that rely on social learning, communication, and knowledge of others' minds (Herrmann *et al.*, 2007). There are also differences in the ways that apes and children solve problems that rely on collaboration. Although chimpanzees are able to solve collaborative tasks effectively (Melis *et al.*, 2006a, 2006b), the means they use to achieve success are different from the means that children use. Chimpanzees do not develop the kinds of joint attention skills that are seen in young children (Tomasello and Carpenter, 2005), and they do not perform well in tasks that build upon this capacity (Wyman and Tomasello, 2007). Chimps coordinate individual goals and their accompanying actions, while "children form shared goals and achieve them through the adoption and reversal of designated roles" (Wyman and Tomasello, 2007: 230).

EMPIRICAL EVIDENCE FOR EMPATHY AND SYMPATHY

Knowledge of others' thoughts, intentions, and desires may give chimps the capacity for empathy. In order to feel sympathy, chimps must also be concerned about the welfare of others. The literature on empathy and sympathy in other primates consists of a number of singular accounts of unusual events, descriptions of several types of common behaviors, and a very small number of systematic analyses and experiments (for a more complete discussion of this body of evidence, see Silk, 2007b).

Conclusions derived from anecdotes are problematic, because they cannot be tested systematically against alternative hypotheses. (To provide a concrete example of this problem, consider the case of Binti Jua. Several years ago, a small child fell into the gorilla enclosure at the Brookfield Zoo in Chicago. Binti Jua, then a young adult female, picked the unconscious child up cradled him to her chest, and eventually turned him over to the zoo staff unharmed. This event was recorded on amateur video, and Binti Jua became an instant celebrity. Some have cited this incident as evidence for empathy and sympathy in apes, arguing that Binti Jua was motivated by compassion and concern for the welfare of the child (Preston and de Waal, 2002). However, other facts need to be considered. Binti Jua was hand-reared by humans, after being rejected by her own mother. Concerned that Binti Jua might become a neglectful mother herself, the zoo staff used operant training methods to guide the development of appropriate maternal skills. One of the things that she was trained to do was to retrieve a doll-like object and bring it to the front of the enclosure, where zoo personnel could inspect it (C. Dimitrios, personal communication). This raises the possibility that Binti Jua's response reflected her training, not her understanding of the child's plight and concern for his welfare.) Compilations of anecdotes suffer from the same problem that plagues all *ad libitum* data collection schemes – they are subject to various sorts of bias (Altmann, 1974; Sarringhaus *et al.*, 2005). For example, observers may be more likely to notice and remember incidents that seem to indicate that monkeys or apes are empathetic or sympathetic about the welfare of others than they are to take note when they seem oblivious and indifferent.

Some forms of common behaviors have been interpreted as evidence of empathy and sympathy. These include wound-cleaning (Boesch, 1992) and consolation (O'Donnell, 1995; de Waal, 1996). Non-human primates often lick and groom others' wounds, and this may play a role in keeping the wounds clean and preventing infection. Boesch (1992) conjectures that wound-cleaning in chimpanzees is based on awareness of the needs of the wounded individual and sympathy for the discomfort that the other is suffering. This conjecture could be correct, but it is also possible that chimps perform this behavior without conscious awareness of others' needs, or because they like the slightly salty taste of blood.

Consolation behavior occurs when bystanders approach, embrace, touch, and groom the victims of aggression, particularly after episodes that include aggressive vocalizations or physical contact (de Waal and Aureli, 1996). Consolation behavior has now been described in several groups of chimpanzees and bonobos (Arnold and Whiten, 2001; Wittig and Boesch, 2003; Palagi *et al.*, 2004). De Waal and Aureli (1996) suggest that chimpanzees console victims of aggression because they empathize with their pain and distress, and are concerned about their welfare. This interpretation implies that consolation will provide an effective means to relieve victims' distress after conflicts. In primates, self-directed behaviors, such as scratching and body shake, are correlated with cortisol levels, and provide an external index of stress levels (Aureli and Smucny, 2000). In a group of captive chimpanzees, consolation did not reduce the rates of self-directed behavior in the victims of aggression (Koski and Sterck, 2007). These data cast doubt on the link between consolation behavior and empathy for the victims of aggression.

Parr (2001) measured the physiological responses of three adult chimpanzees who were shown images of hypodermic needles and dart guns, chimpanzees being darted or injected, and chimpanzees being chased by a veterinarian with a dart gun. Scenes from the home environment (including activity by caretakers, unfamiliar chimpanzees in neutral activities, cage mesh, and transport boxes) were used as control stimuli. Using skin temperature to measure arousal, Parr found that the chimpanzees responded strongly to the images of other chimps being injected and darted, and to images of the dart gun and needles alone. However, the needles and dart guns created as strong a response as the images of chimpanzees being injected or darted. If the chimpanzees' responses were influenced by an understanding of other individuals' feelings or desires, and if they were concerned about the welfare of others, they would have been more strongly affected by the images of other chimpanzees being injected or darted than by the needles or dart guns alone.

SOCIAL PREFERENCES IN PRIMATES

To investigate the nature of social preferences in primates, researchers have recently devised a series of experiments in which animals are presented with opportunities to provide benefits to others at little or no cost to themselves. The choices that they make in these experiments provide insight about their social preferences. The results of this body of work are not fully consistent. Some works suggest that chimps are indifferent to the welfare of other group members, while other works suggest that chimps are motivated to provide benefits to others. Below, I describe these experiments, and then evaluate possible explanations for why the chimps behave differently in different contexts. To help readers keep track of this body of work, a brief synopsis of the experiments is provided in Table 18.1.

Chimps Display Indifference About the Welfare of Other Group Members

My colleagues and I presented chimpanzees from two different captive facilities with the opportunity to provide food rewards to other individuals and to themselves (Silk *et al.*, 2005). To implement their choices, the chimpanzees manipulated experimental

TABLE 18.1 Outline of chimpanzee experimental protocols

		Payoff structure Reward for actor/recipient						
Source	**Location**	**Option 1**		**Option 2**	**Reward type**	**Recipient status**	**Subject ages (mean years)**	**Subject Sex (M, F)**
Silk *et al.*, 2005	Louisiana	1/1	or	1/0	Food	Group member	15	1, 6
	Texas						28	3, 8
Jensen *et al.*, 2006	Leipzig	1/1	or	1/0	Food	Group member	19	2, 9
	Leipzig	0/1	or	0/1	Food	Group member		
	Leipzig	0/1	or	0/1	Food	Group member		
Vonk *et al.*, 2008	Louisiana	1/0	and	0/1	Food	Group member	15	1, 6
	Texas					Group member	28	3, 8
Warneken and Tomasello, 2005	Leipzig	0/1	or	0/0	Non-food	Familiar human	4	1, 2
Warneken *et al.*, 2007	Uganda	0/1	or	0/0	Non-food	Unfamiliar human	10	15, 21
		0/1	or	0/0	Non-food	Group member		
		0/1	or	0/0	Non-food	Group member		

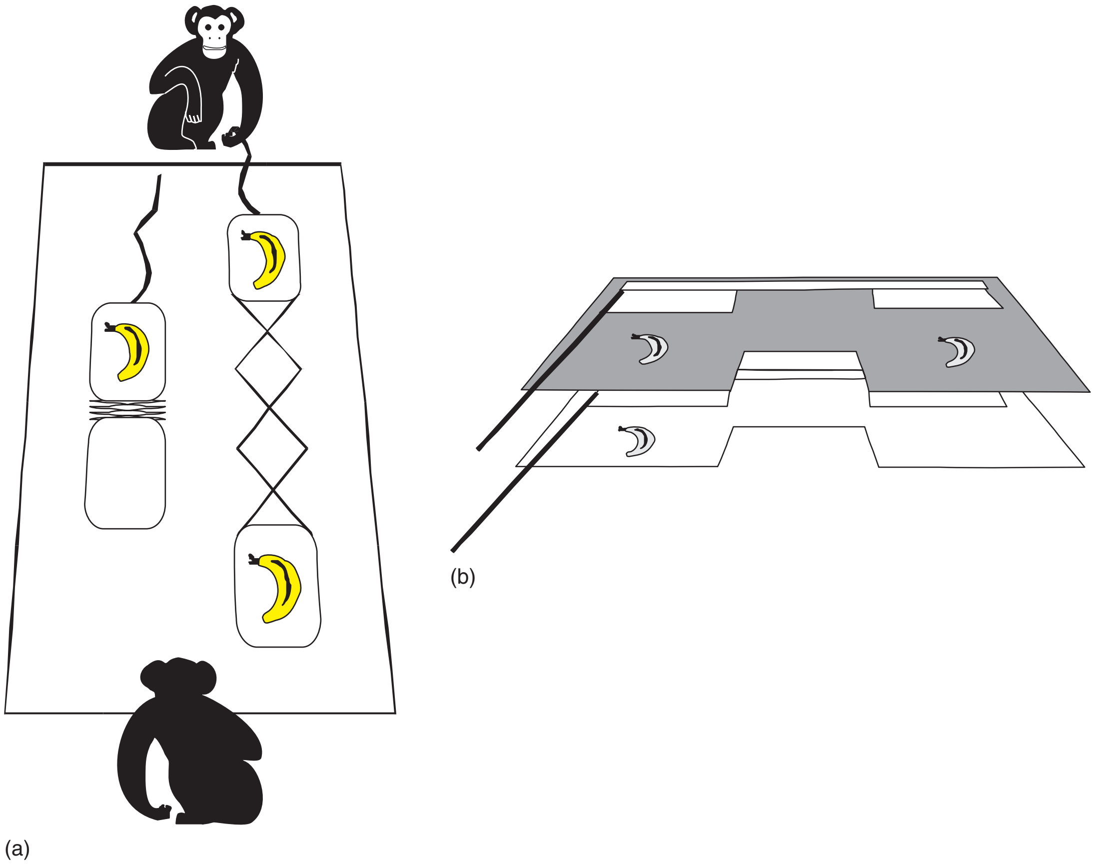


FIGURE 18.3 (a) The actor pulled a rope to expand the device and bring food trays to within reach of its own enclosure and the enclosure on the opposite side. The actor and potential recipient were able to see how the trays were baited, and could also see one another. (b) The actor pulls on one of the two hoses, which sweeps the food forward. When one hose is pulled, the other sweeper is locked in place. Figures drawn by Ruby Boyd.

apparatuses that were baited with food. At one site, the chimps used an expanding device that was positioned between two food trays. When a rope was pulled, the device expanded and one food tray moved to within reach of the actor while the other tray moved to within reach of a chimpanzee (when present) in the opposite enclosure. Two of these devices were placed side by side in a central enclosure (Figure 18.3a). At the other site, the chimpanzees manipulated a two-tier bar pull device. On each tier, a hose was attached to a bar, and when the hose was pulled the bar moved forward and swept food rewards to the front of the platform (Figure 18.3b). One side of each platform was accessible to the actor; the other side was only accessible to the occupant of the adjoining enclosure. At both sites, the location of the potential recipient's food rewards was counterbalanced (right/left, top/bottom) across trials.

In this experiment, the chimps were provided with two options. One option provided identical food rewards to the actor and to the occupant of the other enclosure; the other option provided a food reward only to the actor. These are referred to here as the 1/1 and 1/0 options, respectively (the actor's payoff is given on the left, the recipient's payoff is given on the right). The chimps could choose one of these two options, or do nothing. We realized that chimps might prefer the 1/1 option because they have prepotent biases toward larger numbers of rewards (regardless of the distribution), so a control condition was included in which no potential recipient was present. If individuals are concerned about the welfare of others, we would expect them to prefer the 1/1 option over the 1/0 option, and this preference to be stronger when another individual is present than when the actor is alone. Alternatively, if individuals view potential recipients as rivals or competitors, they may be motivated to deprive them of resources. If so, we would expect them to prefer the 1/0 option over the 1/1 option, and this preference to be stronger when another individual is present than when the actor is alone. Finally, if chimps are indifferent to the welfare of others, we would expect them to choose at random, and their choices not to be affected by the presence of conspecifics.

At both sites, the chimps were as likely to choose the 1/1 option when another chimpanzee was present as when they were alone. Moreover, they were more strongly influenced by the location of the food rewards (left/right, top/bottom) than by the presence of the potential recipient. Based on these findings, we concluded that chimpanzees were indifferent to the welfare of other group members.

Interestingly, cooperatively breeding marmosets showed clear prosocial preferences when they were tested using the same basic protocol that was used in Silk *et al*. (2005). In this experiment, actors were presented with a choice between 0/1 and 0/0, so they did not benefit directly from either choice. The marmosets were significantly more likely to choose the 0/1 option when another marmoset was present than when they were alone (Burkhart *et al*., 2008).

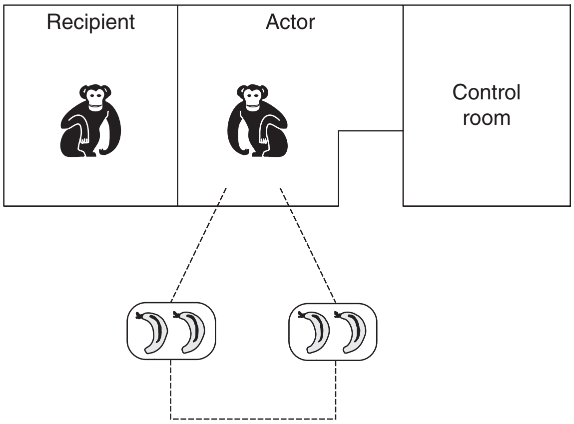


FIGURE 18.4 The actor is confined to the middle room, and can pull a rope that moves the two tables. When one table is pulled forward, the other moves out of reach. The room on the left is occupied by a potential recipient (another chimp from the same group), and the room on the right is always empty. When the door between the actor's room and the potential recipient's room is closed, the actor can only obtain food from its own side of the table. The actor and the potential recipient can see one another, and both can see how the trays are baited. Redrawn from Jensen *et al*. (2006), with permission.

Jensen and colleagues have conducted a different series of experiments to evaluate social preferences in chimpanzees. In this experiment, there were three adjacent enclosures (Jensen *et al*., 2006; Figure 18.4). Two tables were located outside the enclosure, out of the chimps' reach. When one table was pulled forward, the other moved away. When the table on the left side of Figure 18.4 was pulled forward, the inner food cup was accessible to the actor and the outer food cup was accessible to the occupant of the enclosure on the left. When the table on the right side of Figure 18.3 was pulled forward, the inner food cup was accessible to the actor but the outer food cup was not accessible to the occupant of the enclosure on the right. Food cups were positioned on the tables so that only one cup could be reached by the actor. In each trial, the actor was able to move one of the two tables or do nothing, but could not move both tables.

To make sure that the chimps understood that food on the left side was accessible, a set of trials was conducted in which the door between the actor's room and the recipient's room was left open. All four cups were baited, and the chimps were allowed to move only one table. The chimps were significantly more likely to choose the left table (thereby obtaining two food items) than the right table (which provided only one food item), suggesting that they understood that food was accessible from the table on the left, but not from the table on the right. The chimps were tested with members of their social groups with whom they had long-term social relationships.

In test trials, the potential recipient was in the room on the left, and could obtain food if the left-hand table was pulled forward. In control trials, the potential recipient was in the room on the right, and was unable to reach food when the right-hand table was moved forward. The chimps strongly preferred the table on the left. However, they were just as likely to choose the table on the left in the test and in control conditions. Thus, their responses were not affected by whether another individual would profit from their actions. As in the experiments described above, the chimps' were apparently indifferent to the payoffs other chimps obtained.

Jensen and colleagues also tested the chimps' preferences when they did not receive any food themselves. In this experiment, only the outer cups were baited; the actors were able to pull one of the two tables forward, or do nothing. In test trials, the left enclosure was occupied; in control trials, the right enclosure was occupied. The absence of rewards for themselves substantially reduced the chimps' motivation to pull the tables forward. When the inner cups were baited, the chimps pulled one of the tables forward in about 85% of all trials; when the inner cups were empty, the chimps' response rate dropped to about 50%. The actors' preference for the table on the left also declined. When the chimps did make a response, they were as likely to choose the table on the right as the table on the left. Thus, they showed no preference for the option that provided rewards to other individuals.

The same reward distribution was used in the third experiment, but this time the actor had a chance to *prevent* the chimp in the left enclosure from obtaining food. If the actor did nothing, the table on the left would automatically be delivered to the chimp in the enclosure on the left. However, if the actor pulled on a rope, the table on the right would be pulled forward and the recipient would be deprived of food Again, the chimps were most likely to do nothing, and they did not differentiate between the two tables.

In another group of experiments using food rewards, Vonk *et al*. (2008) tested the same chimpanzees that were tested by Silk *et al*. (2006). In this experiment, one choice delivered a reward to the actor, but nothing to the potential recipient (1/0). The other choice delivered an identical reward to the potential recipient, but nothing to the actor (0/1). The actors had the opportunity to choose 1/0 only, 1/0 and 0/1, or 0/1 only, or to do nothing. At one site, the chimps had to use a pole to dislodge a food reward, which rolled down a ramp to the actor or toward the potential recipient (Figure 18.5). At the other site, we used the same two-tiered platform shown in Figure 18.3, but in this case the chimps were able to manipulate both levels. As before, we compared the chimps' responses when another chimp was present in the opposite or adjacent enclosure, with when they were alone.

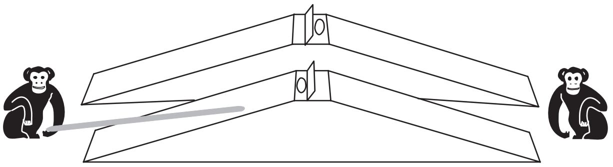


FIGURE 18.5 In this experiment, the actor can use a pole to dislodge a transparent capsule containing a food reward. After it is dislodged, the capsule rolls down the ramp to within reach of the actor or the other chimpanzee (when present). A mesh barrier at the apex of the ramp prevents the food reward on the actor's side from rolling down the ramp toward the potential recipient, and prevents the potential recipient's reward from rolling down the ramp toward the actor.

The chimps were strongly motivated to obtain food for themselves, and did so on virtually every trial in which they made any response. At both sites, the chimps nearly always obtained their own reward first. The chimps sometimes chose the 0/1 option as well, but the presence of a potential recipient did not affect the likelihood that they would do so.

These seven experiments, conducted with four different apparatuses, in three different populations of chimpanzees, generated strikingly similar results. The chimps who participated in these experiments were strongly motivated to obtain rewards for themselves, but did not take advantage of the opportunity to provide food rewards to others at little or no cost to themselves.

It is important to emphasize that in these experiments, the bar for prosocial responses was deliberately set very low. Actors incurred virtually no costs when they behaved prosocially, and they did not have to sacrifice their own rewards to provide rewards to others. This means that other-regarding sentiments did not conflict with selfish motives to obtain rewards, because actors' choices had no effect their own payoffs. In addition, the experiments involved familiar group members, not strangers. Actors might have behaved generously toward group members with whom they cooperated outside the experiment, even if they lacked genuine concern for the welfare of their partners. But the *absence* of prosocial behaviour toward familiar group members implies that actors do not have prosocial preferences about the distribution of food rewards.

Chimps Respond Positively to the Needs of Others

Warneken and Tomasello (2006) showed that young chimpanzees took advantage of opportunities to provide instrumental assistance to familiar human experimenters in some situations. In these experiments, a familiar human trainer tried to accomplish a task, but was unable to do so for various reasons. For each task, a control condition was included in which no help was needed by the experimenter – for example, a book slipped off a stack of books as the adult attempted to place it on top of the stack (experimental), or he placed it next to the stack (control). While human children consistently helped more in the experimental condition than the test condition in four different situations, young chimps only helped consistently in tasks that required retrieving an object that was out of reach. However, they did distinguish between the test and control conditions in several different versions of the reaching/retrieval tasks. Thus, these chimps seemed able to perceive that the human experimenter needed assistance in some situations, and were motivated to provide help.

In a follow-up study, Warneken *et al*. (2007) examined chimps' willingness to extend instrumental help to less familiar humans, and their willingness to help other chimpanzees. In the chimp–human experiments, the subject observed two people struggling over a stick, and one person gaining possession. The victor then placed the stick out of the loser's reach. The loser stretched his arm toward the object (experimental), or simply looked at the object (control). In half of the trials of each condition, the chimps were shown a piece of food which they were given after retrieving the object. Two-thirds of the chimps helped on at least one trial, and they were significantly more likely to help in the experimental condition than in the control condition. However, the presence of a reward did not affect the chimps' behavior.

In a follow-up to this experiment, the chimps were required to exert more physical effort to retrieve the stick; they had to climb 2.5 m up into a passageway. In this experiment, the chimps retrieved the object about half the time. However, they did not distinguish between the experimental and control conditions. Chimps that helped most often in the previous experiment also helped most often in this experiment, suggesting that their failure to distinguish between the test and experimental conditions was "likely due to a carry-over effect from experiment 1 in which subjects had possibly learned that the experimenter wanted the object" (Warneken *et al*., 2007: 1416).

In the third and most compelling experiment in this study, chimps were given an opportunity to provide help to other chimps. In this experiment, the doors to two rooms were fastened by chains held in place by a peg (Figure 18.6). When the peg was removed, the door could be opened. The subject was confined to the room with the pegs. The subject could then remove a peg and release the door on the right side, but could not reach the peg on the left. The chimp in the middle room could not reach the pegs. The door to the room on the left could not be opened by either chimpanzee. In experimental trials, food was placed in the room on the right. In control trials, the food was placed in the room on the left. Nine chimpanzees served as subjects in this experiment, and three unrelated chimps served as recipients.

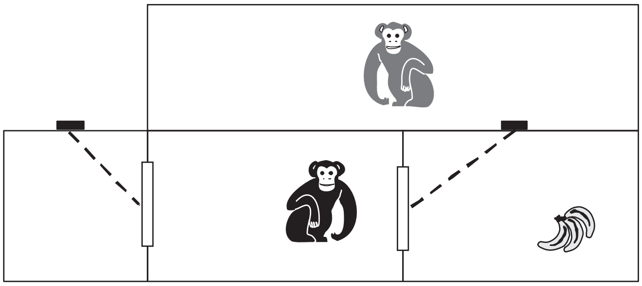


FIGURE 18.6 The actor occupies the room at the top of this diagram. The potential beneficiary occupies the middle room, which is connected to the other two rooms by doors. In order for the potential beneficiary to open these doors, a peg must be released. The actor can reach the peg that releases the door to the room on the right, but cannot reach the peg that releases the door to the room on the left. Food rewards are placed in one of the two connecting rooms, and are visible to the potential beneficiary but not to the actor. Redrawn from Warneken *et al*. (2007), with permission.

The recipients approached the door to the room on the right in every experimental trial, but oriented toward the door to the room on the left only about half the time in control trials. The actors were significantly more likely to release the door in the experimental condition than in the control condition; the difference in the chimps' responses in the experimental and control conditions increased over the course of the experiment. Actors might have been helpful because they expected to share in the rewards obtained. However, the actors never begged for food after the door to the room on the right was opened, and never received food. Thus, these results suggest that the actors were responsive to the needs and desires of other chimps.

Reconciling the Results

These two bodies of experimental work lead to seemingly incompatible conclusions about social preferences in chimpanzees. No serious methodological flaws have been detected in any of the experiments, and both bodies of work are based on multiple experiments that generate internally consistent results. In all of these experiments, appropriate controls for things like location preferences, prepotent biases, and dominance rank were implemented in the design of the experiment or the statistical analyses. Therefore, it seems that we must proceed under the assumption that the results of these experiments are reliable and robust. Below, I examine several possible explanations for the discrepancy in the results obtained by Silk/Jensen/Vonk and by Warneken and colleagues.

1. *The chimps do not display prosocial biases because they do not understand the way the apparatuses work*

It is possible that the Silk experiments did not elicit prosocial responses because the chimps simply did not understand how the experimental apparatuses worked, and did not realize that they could use the apparatus to deliver rewards to other individuals. This criticism implicitly assumes that chimps have prosocial preferences, and the experiments failed to elicit their true preferences.

This objection cannot explain the results obtained by Jensen and his colleagues, because they allowed the chimps access to the enclosure on the left side of the apparatus. When they were allowed to enter the room on the left, the chimps showed clear preferences for the table that delivered rewards that could be recovered by the occupant of the left enclosure over the table that delivered rewards to the other room. Thus, there is little doubt that the chimps knew how the apparatus used in these experiments worked.

My colleagues and I did not explicitly test the chimps' understanding of the apparatuses that we used in our experiments by allowing them to obtain rewards from the recipients' enclosure. Instead, we relied on preliminary tests which demonstrated that the actors (a) understood that they were able to obtain rewards from one side/level of the apparatus, but not the other; (b) were able to counteract strong location biases when it affected their payoffs; and (c) were attending to the distribution of rewards. At the site where the expanding apparatus was used, all of the chimps participated in the experiments as actors and recipients, giving them a chance to experience both roles. The two-tiered platform device that was used at the other is quite similar to one used in many other studies of cooperation in monkeys and apes, and was not expected to present cognitive difficulties for the chimps in our experiments. Moreover, in the experiments conducted by Vonk *et al*. (2008), the likelihood of choosing the other reward declined across trials, suggesting that the chimps understood that they would not obtain the reward delivered to the other enclosure. Finally, the positive results obtained with marmosets, whose cognitive abilities are considerably more limited than the cognitive abilities of apes, suggest that it is unlikely that the chimps did not understand the way that the apparatus worked.

2. *Tasks involving food rewards do not produce prosocial responses in chimpanzees because chimps perceive others as competitors over access to limited quantities of food* **(Warneken and Tomasello, 2006; Warneken et al., 2007)**

This explanation focuses on the selective pressures that shape the evolution of social preferences. Chimpanzees live in a more individualistic world than callitrichids, and this corresponds to differences in their performance on the prosocial task (Burkardt *et al*., 2008). But it is not clear that this explanation fits the chimps' performance on the tasks devised by Silk/Jensen/Vonk. These tasks all involved small food rewards, such as one slice of banana. If chimps perceived other chimps as competitors and rivals over access to limited quantities of food, then they would be expected systematically to deny others access to food, not to ignore them. The chimps didn't behave this way. For example, in the Silk/Jensen experiments, the chimps did not systematically prefer the 1/0 option over the 1/1 option. Instead, they behaved as if they were indifferent to the presence of others.

3. *The prospect of obtaining food for themselves might have made the chimps oblivious to the needs and desires of other chimpanzees*

When chimps were faced with a choice between 1/1 and 1/0 options, they may have been absorbed by their own food rewards and ignored the effects of their choices on others. However, the responses of the chimps were essentially the same when they did not obtain any food rewards themselves. Recall that Jensen and his colleagues conducted two experiments in which actors were unable to obtain food for themselves. The chimps were considerably less motivated to respond in this situation, suggesting that they understood that they would not obtain rewards themselves. Nonetheless, even when the prospect of obtaining food was eliminated, the chimps did not show prosocial preferences. Similarly, in Vonk *et al*. (2008), there was a long latency between dislodging the two rewards. After dislodging and consuming their own rewards, the chimps had considerable time to attend to the needs of their partners.

4. *Chimps respond to direct requests for help, but do not take advantage of opportunities to provide unsolicited assistance*

Prosocial responses might not have been observed in the Silk/Jensen experiments because the actors "were preoccupied with retrieving food for themselves, and the recipient did nothing to indicate any need for help" (Warneken *et al.*, 2007). However, this is not entirely correct. Potential recipients could and did use gestures to attract the actor's attention and request food in some trials in the Silk/Vonk experiments. As noted earlier, Vonk and colleagues showed that potential recipients gestured before actors had delivered food rewards to them in over half the trials. These gestures ought to have attracted the attention of the actors, and alerted them to the recipient's presence and desires. However, gestures by potential recipients had no significant impact on the likelihood of receiving rewards.

We have also reviewed the videotaped records of the behavior of potential recipients in the experiments presented in our earlier experiments (Silk *et al.*, unpublished data). During these trials, we coded the recipient's behavior before actors had had a chance to make a choice between the 1/1 and 1/0 options. Recipients gestured toward the actor and food trays before the actor was able to make a choice in about one-third of all trials. Recipients nearly always gestured toward the 1/1 side of the apparatus. When recipients gestured before a choice was made, actors chose the 1/1 option 60% of the time. When recipients did not gesture before a choice was made, actors chose the 1/1 option 56% of the time. Again, gestures by potential recipients did not have a consistent impact on the actor's behavior.

5. *Differences in the chimps' responses reflect variation in their socialization experiences, rearing histories, age, or sex*

Efforts have been made to assess the effects of age and sex on variation in performance within experiments (see, for example, Warneken *et al.*, 2007), but it is also worth considering the possibility that demographic factors, socialization experiences, or rearing histories generated variation in performance across experiments. All of the subjects were members of bisexual social groups and were maintained in captive facilities. The chimps tested by Warneken *et al.* (2007) were born in the wild, but subsequently orphaned and confiscated by wildlife authorities. They now live together in a well-maintained 95-acre sanctuary in Uganda.

The chimps that participated in the Silk/Jensen/Vonk experiments were substantially older than most of the chimps tested by Warneken and his colleagues (Table 18.1). Chimpanzees reach sexual maturity in their teens (Muller and Mitani, 2005), so the majority of subjects tested by Silk/Jensen/Vonk were adults, while the majority of chimps tested by Warneken and his colleagues were juveniles or adolescents. We know nothing about the ontogenetic development of cooperative behavior in chimpanzees or other primates, leaving open the possibility that differences in performance on these experimental tasks reflects differences in the maturity of the subjects.

CONCLUSIONS

It is more difficult to demonstrate that help is based on prosocial preferences than to demonstrate the absence of prosocial preferences in a particular experimental setting. This is because the existence of altruism is not necessarily evidence for prosocial preferences. Altruistic behavior occurs in species that lack a well-developed theory of mind and the capacity for empathy. Altruism even occurs in social microbes that don't have brains at all (Robinson *et al.*, 2005). This means that the motives and social preferences that underlie altruistic behavior must vary across species, and may differ among humans and other primates.

The experiments conducted by Warneken *et al.* (2007) show that chimps are sometimes willing to provide instrumental assistance to others, but the motives that underlie behavior in these experiments are not clear. Actors might provide instrumental assistance toward group members with whom they have cooperative relationships based on reciprocity outside the experiment, even if they lack genuine concern for the welfare of their partners. In contrast, chimps did not demonstrate prosocial preferences in any of the experiments that involved food rewards. If chimps do have other-regarding preferences, these preferences are clearly muted among adult chimps when food is present.

Although the discrepancy between the results obtained in experiments that were designed to examine chimps' social preferences may seem like an unfortunate complication, it provides us with an important opportunity. If we can work out the reasons that the results of these experiments differ, we may gain deeper insights about the complexity and nature of chimps' social preferences. For example, the presence of food rewards might compete with selfish motives about food. If that is the case, then the chimps who were indifferent in the Silk/Jensen/Vonk experiments might behave prosocially in the protocol used by Warneken and his colleagues, and *vice versa*. Moreover, it would be straightforward to modify the experimental procedures used by Warneken and colleagues to examine chimps' willingness to retrieve food items for humans or other chimps. To move the discussion forward, we need to make systematic efforts to explain why individuals make different choices in different settings.

Acknowledgments

My work on prosocial preferences was conducted in collaboration with Joe Henrich, Sarah Brosnan, Jennifer Vonk, and Daniel Povinelli, and was funded by a grant from the MacArthur Preferences Network. I thank Ruby Boyd for preparing the figures, and Robert Boyd for comments on an earlier draft of this paper.

References

Altmann, J. (1974). Observational study of behavior: sampling methods. *Behaviour* 49, 227–267.

Arnold, K. and Whiten, A. (2001). Post-conflict behaviour of wild chimpanzees (*Pan troglodytes schweinfurthii*) in the Budongo forest, Uganda. *Behaviour* 138, 649–690.

Aureli, F. and Smucny, D. (2000). The role of emotion in conflict and conflict resolution. In: F. Aureli and F.B.M. de Waal (eds), *Natural Conflict Resolution*. Berkeley, CA: University of California Press, pp. 199–224.

Axelrod, R. and Hamilton, W.D. (1981). The evolution of cooperation. *Science* 211, 1390–1396.

Bales, K., Dietz, J., Baker, A. *et al*. (2000). The effects of allocaregivers on fitness of infants and parents in callitrichid primates. *Folia Primatol.* 71, 27–38.

Batson, C.D. (1991). *The Altruism Question: Toward a Social–Psychological Answer*. Hillsdale, NJ: Erlbaum.

Boyd, R. and Richerson, P.J. (2005). *Not by Genes Alone*. Chicago, IL: University of Chicago Press.

Boyd, R., Gintis, H., Bowles, S., and Richerson, P.J. (2003). The evolution of altruistic punishment. *Proc. Natl Acad. Sci. USA* 100, 3531–3535.

Boesch, C. (1992). New elements of a theory of mind in wild chimpanzees. *Behav. Brain Sci.* 15, 149–150.

Burkhart, J., Fehr, E., Efferson, C., and van Schaik, C. (2008). Other-regarding preferences in a nonhuman primate: common marmosets provision food altruistically. *Proc. Natl Acad. Sci.* USA 104, 19762–19766.

Call, J. (2007). Social knowledge in primates. In: R.I.M. Dunbar and L. Barrett (eds), *Oxford Handbook of Evolutionary Psychology*. Oxford: Oxford University Press, pp. 71–81.

Chapais, B. and Berman, C.M. (2004). *Kinship and Behavior in Primates*. Oxford: Oxford University Press.

Cheney, D.L. and Seyfarth, R.M. (2007). *Baboon Metaphysics: The Evolution of a Social Mind*. Chicago, IL: University of Chicago Press.

Clutton-Brock, T.H. and Parker, G.A. (1995). Punishment in animal societies. *Nature* 373, 209–216.

Cords, M. (2002). Friendship among adult female blue monkeys (*Cercopithecus mitis)*. *Behaviour* 139, 291–314.

Crozier, R.H. and Pamilo, P. (1996). *Evolution of Social Insect Colonies. Sex Allocation and Kin Selection*. Oxford: Oxford University Press.

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection*. London: John Murray.

de Waal, F.B.M. (1996). *Good Natured*. Cambridge, MA: Harvard University Press.

de Waal, F.B.M. and Aureli, F. (1996). Consolation, reconciliation, and a possible cognitive difference between macaques and chimpanzees. In: A.E. Russon, K.A. Bard, and S.T. Parker (eds), *Reaching into Thought*. Cambridge: Cambridge University Press, pp. 80–110.

Dugatkin, L.A. (1997). *Cooperation among Animals*. Oxford: Oxford University Press.

Dugatkin, L.A. and Reeve, H.K. (1994). Behavioral ecology and levels of selection: dissolving the group selection controversy. *Adv. Study Behav.* 23, 101–133.

Duffy, K.G., Wrangham, R.W., and Silk, J.B. (2007). Male chimpanzees exchange political support for mating opportunities. *Curr. Biol.* 21, R586–587.

Fehr, E. and Fischbacher, U. (2003). The nature of human altruism. *Nature* 425, 785–791.

Flack, J.C., de Waal, F.B.M., and Krakauer, D.C. (2005). Social structure, robustness, and policing cost in a cognitively sophisticated species. *Am. Naturalist,* 165, E126–139.

Flombaum, J.I. and Santos, L.R. (2005). Rhesus monkeys attribute perceptions to others. *Curr. Biol.* 15, 447–452.

Ghazanfar, A.A. and Santos, L.R. (2004). Primate brains in the wild: the sensory bases for social interactions. *Nat. Rev. Neurosci.* 5, 603–616.

Gintis, H., Bowles, S., Boyd, R., and Fehr, E. (2007). Explaining altruistic behaviour in humans. In: R.I.M. Dunbar and L. Barrett (eds), *Oxford Handbook of Evolutionary Psychology*. Oxford: Oxford University Press, pp. 605–619.

Goodall, J. (1986). *The Chimpanzees of Gombe: Patterns of Behaviour*. Cambridge, MA: Belknap Press.

Gros-Louis, J. (2004). The function of food-associated calls in white-faced capuchin monkeys, *Cebus capucinus*, from the perspective of the signaller. *Animal Behav.* 67, 431–440.

Haig, D. (1999). What is a marmoset? *Am. J. Primatol.* 49, 285–296.

Hamilton, W.D. (1964). The genetical evolution of social behavior. I and II. *J. Theor. Biol.* 7, 1–52.

Hammerstein, P. (2003). Why is reciprocity so rare in social animals? A Protestant appeal. In: P. Hammerstein (ed.), *Genetic and Cultural Evolution of Cooperation*. Cambridge, MA: MIT Press, pp. 83–94.

Hare, B., Call, J., Agnetta, B., and Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behav.* 59, 771–785.

Hare, B., Call, J., and Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behav.* 61, 139–151.

Hare, B., Addessi, E., Call, J. *et al*. (2003). Do capuchin monkeys, *Cebus apella*, know what conspecifics do and do not see? *Animal Behav.* 65, 131–142.

Hauser, M.D. and Marler, P. (1993a). Food-associated calls in rhesus macaques (*Macaca mulatta*): I. Socioecological factors. *Behav. Ecology* 4, 194–205.

Hauser, M.D. and Marler, P. (1993b). Food-associated calls in rhesus macaques (*Macaca mulatta*): II. Costs and benefits of call production and suppression. *Behav. Ecology,* 4, 206–212.

Hemelrijk, C.K. (1994). Support for being groomed in long-tailed macaques. *Macaca fasicularis. Animal Behav.* 48, 479–481.

Henrich, J., McElreath, R., Barr, A. *et al*. (2006). Costly punishment across human societies. *Science* 312, 1767–1770.

Herrmann, E., Call, J., Hernández-Lloreda, M.V. *et al*. (2007). Humans have evolved specialized skills in social cognition: the cultural intelligence hypothesis. *Science* 317, 1360–1366.

Jensen, K., Hare, B., Call, J., and Tomasello, M. (2006). What's in it for me? Self-regard precludes altruism and spite in chimpanzees. *Proc. R. Soc. Biol.* 273, 1013–1021.

Jensen, K., Call, J., and Tomasello, M. (2007a). Chimpanzees are vengeful but not spiteful. *Proc. Natl Acad. Sci. USA* 104, 13046–13050.

Jensen, K., Call, J., and Tomasello, M. (2007b). Chimpanzees are rational maximizers in an Ultimatum Game. *Science* 318, 107–109.

Kappeler, P.M. and van Schaik, C. (2006). *Cooperation in Primates and Humans: Mechanisms and Evolution*. Berlin: Springer.

Kapsalis, E. (2004). Matrilineal kinship and primate behavior. In: B. Chapais and C.M. Berman (eds), *Kinship and Behavior in Primates*. Oxford: Oxford University Press, pp. 153–176.

Karin-D'Arcy, M.R. and Povinelli, D.J. (2002). Do chimpanzees know what each other see? A closer look. *Intl J. Comp. Psych.* 15, 21–54.

Koski, S.E. and Sterck, E.H.M. (2007). Triadic postconflict affiliation in captive chimpanzees: does consolation console? *nimal Behav.* 73, 133–142.

Langergraber, K.E., Mitani, J.C., and Vigilant, L. (2007). The limited impact of kinship on cooperation in wild chimpanzees. *Proc. Natl Acad. Sci USA* 104, 7787–7790.

McElreath, R. and Boyd, R. (2007). *Mathematical Models of Social Evolution: A Guide for the Perplexed*. Chicago, IL: University of Chicago Press.

Melis, A.P., Hare, B., and Tomasello, M. (2006a). Engineering cooperation in chimpanzees: social tolerance constraints on cooperation. *Animal Behav.* 72, 275–286.

Melis, A.P., Hare, B., and Tomasello, M. (2006b). Chimpanzees recruit the best collaborators. *Science* 311, 1297–1300.

Muller, M.N. and Mitani, J.C. (2005). Conflict and cooperation in wild chimpanzees. *Adv. Study Behav.* 35, 275–331.

Nishida, T., Hosaka, K., Nakamura, M., and Hamai, M. (1995). A within-group gang attack on a young adult male chimpanzee: ostracism of an ill-mannered member? *Primates* 36, 207–211.

Noë, R. (2005). Digging for the roots of trading. In: P.M. Kappeler and C.P. van Schaik (eds), *Cooperation in Primates and Humans: Mechanisms and Evolution*. Berlin: Springer, pp. 223–251.

O'Donnell, S. (1995). Empathy in chimpanzees: evidence for theory of mind. *Primates* 36, 397–410.

Palagi, E., Paoli, T., and Borgognini Tarli, S. (2004). Reconciliation and consolation in captive bonobos (*Pan paniscus*). *Am. J. Primatol.* 62, 15–30.

Palombit, R.A. (1994). Extra-pair copulations in a monogamous ape. *Animal Behav.* 47, 721–723.

Panchanathan, K. and Boyd, R. (2005). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* 432, 499–502.

Parr, L. (2001). Cognitive and physiological markers of emotional awareness in chimpanzees (*Pan trogloydytes*). *Animal Cogn.* 4, 223–229.

Preston, S.D. and de Waal, F.B.M. (2002). Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* 25, 1–72.

Pusey, A.E. (2004). Inbreeding avoidance in primates. In: A.P. Wolf and W.H. Durham (eds), *Incest, Inbreeding and the Incest Taboo*. Stanford, CA: Stanford University Press, pp. 61–75.

Ratnieks, F.L.W. (1988). Reproductive harmony via mutual policing by workers in eusocial Hymenoptera. *Am. Naturalist,* 132, 217–236.

Ratnieks, F.L.W. and Wenseleer, T. (2005). Policing insect societies. *Science* 307, 54–56.

Reeve, H.K. and Keller, L. (1997). Reproductive bribing and policing as evolutionary mechanisms for the suppression of within-group selfishness. *Am. Naturalist* 150, S42–S58.

Robinson, G.E., Grozinger, C.M., and Whitfield, C.W. (2005). Sociogenomics: social life in molecular terms. *Nat. Rev. Genetics* 6, 257–270.

Ross, C.N., French, J.A., and Orti, G. (2007). Germ-line chimerism and paternal care in marmosets (*Callithrix kuhlii*). *PNAS* 104, 6278–6282.

Santos, L.R., Nissen, A.G., and Ferrugia, J.A. (2006). Rhesus monkeys, *Macaca mulatta*, know what others can and cannot hear. *Animal Behav.* 71, 1175–1181.

Sarringhaus, L.A., McGrew, W.C., and Marchant, L.F. (2005). Misuse of anecdotes in primatology: lessons from citation analysis. *Am. J. Primatol.* 65, 283–288.

Seyfarth, R.M. and Cheney, D.L. (1984). Grooming, alliances, and reciprocal altruism in vervet monkeys. *Nature* 308, 541–543.

Silk, J.B. (2005). The evolution of cooperation in primate groups. In: H. Gintis, S. Bowles, R. Boyd, and E. Fehr (eds), *Moral Sentiments and Material Interests: On the Foundations of Cooperation in Economic Life*. Cambridge, MA: MIT Press, pp. 43–73.

Silk, J.B. (2007a). The strategic dynamics of cooperation in primate groups. *Adv. Study Behav.* 37, 1–42.

Silk, J.B. (2007b). Empathy, sympathy, and prosocial preferences in primates. In: R.I.M. Dunbar and L. Barrett (eds), *Oxford Handbook of Evolutionary Psychology*. Oxford: Oxford University Press, pp. 115–126.

Silk, J.B., Alberts, S.C., and Altmann, J. (2004). Patterns of coalition formation by adult female baboons in Amboseli, Kenya. *Animal Behav.* 67, 573–582.

Silk, J.B., Brosnan, S.F., Vonk, J. *et al*. (2005). Chimpanzees are indifferent to the welfare of other group members. *Nature* 435, 1357–1359.

Silk, J.B., Altmann, J., and Alberts, S.C. (2006a). Social relationships among adult female baboons (*Papio cynocephalus*) I. Variation in the strength of social bonds. *Behav. Ecol. Sociobiol.* 61, 183–195.

Silk, J.B., Alberts, S.C., and Altmann, J. (2006b). Social relationships among adult female baboons (*Papio cynocephalus*) II: Variation in the quality and stability of social bonds. *Behav. Ecol. Sociobiol.* 61, 197–204.

Stammbach, E. (1987). Desert, forest, and montane baboons: multilevel societies. In: B.B. Smuts, D.L. Cheney, R.M. Seyfarth *et al*. (eds), *Primate Societies*. Chicago, IL: University of Chicago Press, pp. 112–120.

Tomasello, M. and Call, J. (1997). *Primate Cognition*. Oxford: Oxford University Press.

Tomasello, M. and Carpenter, M. (2005). The emergence of social cognition in three young chimpanzees. *Monographs Soc. Res. Child Dev.* 70, 1–152.

Trivers, R.L. (1971). The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57.

Vonk, J., Brosnan, S.F., Silk, J.B. *et al*. (2008). Chimpanzees do not take advantage of very low cost opportunities to deliver food to unrelated group members. *Animal Behav*, (in press).

Warneken, F. and Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science* 311, 1301–1303.

Warneken, F., Hare, B., Melis, A.P. *et al*. (2007). Spontaneous altruism by chimpanzees and young children. *PLoS Biol.* 5, 1414–1420.

Widdig, A. (2007). Paternal kin discrimination: the evidence and likely mechanisms. *Biol. Rev.* 82, 319–334.

Williams, J.M., Liu, H., and Pusey, A.E. (2002). Costs and benefits of grouping for female chimpanzees at Gombe. In: C. Boesch, G. Hohmann, and L. Marchant (eds), *Behavioural Diversity in Chimpanzees and Bonobos*. Cambridge: Cambridge University Press, pp. 192–203.

Wilson, D.S. and Sober, E. (1994). Reintroducing group selection to the human behavioral sciences. *Behav. Brain Sci.* 17, 585–654.

Wittig, R. and Boesch, C. (2003). "Decision-making" in conflicts of wild chimpanzees (*Pan troglodytes*): an extension of the relational model. *Behav. Ecol. Sociobiol.* 54, 491–504.

Wyman, E. and Tomasello, H. (2007). The ontogenetic origins of cooperation. In: R.I.M. Dunbar and L. Barrett (eds), *Oxford Handbook of Evolutionary Psychology*. Oxford: Oxford University Press, pp. 227–236.

CHAPTER

19

Responses to Inequity in Non-human Primates

Sarah F. Brosnan

OUTLINE

INTRODUCTION

Humans are very sensitive to inequity. We respond badly when we are treated unfairly, we give more than the minimum amount required in experimental games, and we frequently punish in situations in which another individual behaves non-cooperatively (Kahneman *et al.*, 1986; Zizzo and Oswald, 2001; Fehr and Rockenbach, 2003). These inequity averse responses are seen across a wide variety of cultures (Henrich *et al.*, 2001), and vary significantly depending upon the quality of the relationship between the individuals involved (Clark and Grote, 2003). They have recently been linked to emotional, as well as rational, processes, both at the level of behavior and at the level of neural activity (Frank, 1988; Sanfey *et al.*, 2003).

About a decade ago, it was proposed that negative reactions to inequity might be beneficial as a mechanism for stabilizing cooperation in human societies (Fehr and Schmidt, 1999). Since then, this approach has been applied in human research, and is beginning to be applied in non-human studies as well. This chapter provides a brief background on inequity aversion theory and research in humans, and then discusses the relevant research in non-human primates. It closes by comparing the two responses and discussing the implications that the non-human data have for the evolution of the inequity response.

INEQUITY AVERSION

Although traditionally economics has used the simplifying assumption that humans are rational, maximizing their absolute gains in a directed way, in the latter half of the twentieth century the work done on experimental games (among other things) began to make it clear that this was not always the case. Either the games weren't tapping in to human economic choice behavior very well, or humans weren't always all that rational with regard to the maximization of their monetary wealth. Some economists began to develop new theories to explain these deviations from expected behavior in a wide variety of economic games. These theories typically incorporated other variables (often social preferences) in humans' calculations of utility.

Among these is a class of models incorporating dislike of inequity into the human utility function, a concept already well-known from research in psychology (Walster *et al.*, 1978). The two primary models, Fehr and Schmidt's (1999) Inequity Aversion (IA) model and Bolton and Ockenfels's (2000) Theory of Equity, Reciprocity, and Competition (ERC) are based on this principle (see Box 19.1 for more details). The models predict that individuals dislike inequity, and therefore players get utility from minimizing the difference in payoffs between themselves and others (i.e. equalizing their relative gains). Fehr and Schmidt's model explicitly included parameters for disadvantageous inequity aversion (the α parameter), inequity which is detrimental to the self, and advantageous inequity aversion (the β parameter), which is inequity advantageous to the self. Together these explain why, for instance, responders in the ultimatum game (see Box 19.2 for details on selected economic games) frequently turn down positive, but relatively much smaller, offers and why proposers in the dictator game ever give anything at all. Aversion to inequity (or an egalitarian motive) has also been argued to explain behavior in random income games (in these games, each player is initially allocated a random amount of money which differs from the allocation of other players), in which individuals will pay to reduce wealthier individuals' incomes or supplement poorer individuals' incomes, even though distribution is random and no intention (positive or negative) was involved (Dawes *et al.*, 2007).

Another thing that these models have in common is the hypothesis that an aversion to inequity can stabilize cooperation. Individually, a sense of inequity is a useful rule of thumb for determining which other players or group members make the best cooperation partners (i.e., cooperate with individuals who share the payoffs). Within a group, acting to rectify inequity can advertise one's commitment to equity and

BOX 19.1

THE FEHR-SCHMIDT MODEL OF INEQUITY AVERSION

Fehr and Schmidt's model of inequity aversion assumes that individuals are interested not only in their own, but also in others' gains. Hence, their model divides inequity into "disadvantageous" inequity, or that which harms the actor, and "advantageous" inequity, or that which harms another. Fehr and Schmidt's (1999) model differs from a similar model, the Equality, Reciprocity, and Competition (ERC) model (Bolton and Ockenfels, 2000) in its focus on absolute (rather than relative) differences in payoffs and its focus on comparisons among players (rather than summations of payoffs; Camerer, 2003). In most comparisons, the Fehr/Schmidt Inequity Aversion model appears to explain behavior better than does the ERC model.

Mathematically, for the players i and k, player i's utility function for the social allocation $X = \{x_1, x_2, \ldots x_n\}$ can be described as:

$U_i(X) = x_i - \alpha/(n-1)\, \Sigma \max(x_k - x_i, 0) - \beta/(n-1)\, \Sigma \max(x_i - x_k, 0)$, where $k \neq i$ for the purposes of the summation. In this formulation, α represents "envy," or disadvantageous inequity aversion, and β represents "guilt," or advantageous inequity aversion. We can also assume that $0 = \beta_i < 1$ and $\beta_i = \alpha_i$, or that disadvantageous inequity aversion is stronger than advantageous inequity aversion. For the ultimatum game (see Box 19.2), for example, if player i is the responder, she should reject any offer less than $\alpha_{i/1+2}\alpha_i$, and the proposer (player k) should make offers dependent upon her β_k, or guilt. For more details, see Fehr and Schmidt, 1999, and Camerer, 2003 and Chapter 15.

BOX 19.2

GAMES IN EXPERIMENTAL ECONOMICS

Experimental economics utilizes a number of "games" which can be used to elucidate subjects' decision making strategies and social preferences in controlled situations. For more details on these games, or additional games, see Chapter 5 of this volume, or Camerer 2003.

The Ultimatum Game

Invented by Werner Güth (Güth *et al*., 1982), the ultimatum game (UG) is a bargaining game. The first player, or proposer, is given a sum of money and told to divide it between herself and a partner (typically anonymous), with the knowledge that the second player, or responder, can either accept the division and both will be rewarded as the proposer indicated, or can refuse the division, in which case neither player receives any money.

Although the most "rational" action is for the proposer to offer the least amount possible and for the responder to accept it (which is a gain for the responder), the game rarely goes this way in practice. The median offer by the proposer in most studies is 50% (the average is about 40%), and responders frequently refuse offers of less than this. This has been interpreted as individuals having a "taste" for fairness (see Box 19.1). Many social or situational factors will affect player's responses in this game.

A recent study has found that chimpanzees act in the "rational" way – that is, responders will accept any offer (Jensen *et al*., 2007). However, rationality also assumes that responders will only accept positive offers, and in this study chimpanzee responders accepted the payoff distribution 56% of the time when they got nothing at all, indicating that some or all of them failed to understand the contingencies of the game.

UG Variation: The Dictator Game

The dictator game (DG) is a variation on the UG which evaluates the fairness goals of the proposer. The DG is not really a game, as the proposer makes a unilateral decision about how to divide the money and the responder makes no decision and cannot affect the outcome. Not surprisingly, proposer divisions are much more lopsided in the DG than in the UG, with much less of the pot offered to the responder.

UG Variation: The Impunity Game

The impunity game (IG) is another variation on the UG in which the responder has only limited recourse. After the proposer makes a division, the responder can reject, but her rejection affects only her – and not the proposer's – payoff. Thus, the rational response is for proposers to offer the least amount possible and for responders to accept it, as in the DG (Bolton and Zwick, 1995).

However, responders often refuse their winnings, both when the proposer will know the responder's actions and when the proposer is ignorant (i.e., the proposer believes them to be playing a DG; Yamagishi, 2007). While this is unexpected from a perspective of rationality, as this response leads to increased (rather than decreased) inequity, it may indicate that people's responses serve not only to equalize outcomes but also to send a signal to both their partners and themselves. Such signals could constitute commitment devices (Frank, 1988) which inform others of the player's refusal to participate in outcomes which do not have fair outcomes, increasing the player's long-term gains in cooperative interactions. Similar responses are seen in non-human primates (Brosnan and de Waal, 2003, Brosnan *et al*., 2005).

The Prisoner's Dilemma Game

The prisoner's dilemma game, a model of decision making when interests do not intersect, allows two individuals to each make a decision to work together (cooperate) or not (defect). The payoff matrix is as follows (player 1 is indicated in bold, player 2 in italics), where Reward is the reward for mutual cooperation, Punishment is the payoff for mutual defection, Temptation is the payoff for defecting, and Sucker is the "Sucker's payoff" for cooperating.

	Cooperate	**Defect**
Cooperate	*Reward*, **Reward**	*Sucker*, **Temptation**
Defect	*Temptation*, **Sucker**	*Punishment*, **Punishment**

In the standard game, Temptation > Reward > Punishment > Sucker's payoff. Thus, in a non-repeated (one-shot) game, the optimum strategy is to defect no matter which decision one's partner makes, as the payoffs are higher (T > R, P > S). However, in a repeated (or iterated) game, the best payoff comes from cooperation, because in the long run both players maximize their payoffs with mutual cooperation, rather than mutual defection (R > P). This, of course, assumes two players who will not continue to cooperate in the absence of cooperation from a partner. In an iterated game, the best strategy is some variation on "tit-for-tat," in which each player chooses the option her partner chose on the previous round (developed by Anatol Rapaport; see Axelrod, 1984, for more details).

to one's cooperative partners (Frank, 1988), and can reduce the frustration of the less well-endowed individuals. If, in fact, inequity aversion is a mechanism for stabilizing cooperation, it is reasonable to assume that such behavior may exist in other cooperative species, including our non-human primate relatives.

Aversion to inequity implies, of course, a reaction to both over- and under-compensation, although these are likely not treated the same by most people. While it is often found that people prefer equity to any sort of inequity, advantageous inequity is typically preferred to disadvantageous inequity (Loewenstein *et al.*, 1989). Moreover, sometimes inequity, such as that favoring an *in-group* (one's own group), is even preferred (Diekmann *et al.*, 1997; Bernhard *et al.*, 2006). Furthermore, people tend to rectify inequity through psychological balancing mechanisms – justifying why they deserve a greater share – rather than material compensation (Walster *et al.*, 1978), and will usually choose to ignore information which could lead to a more fair outcome at a cost to the self (Dana *et al.*, 2003). Thus, it appears, humans are interested in fairness as an ideal, but self-interest still comes first in most situations.

Inequity aversion may not be the only explanation for these behaviors. A second set of models, the reciprocity models, are invoked to explain similar behavior patterns as do the inequity aversion models. These models incorporate intentions, and assume that much behavior can be explained by individuals' attempts to reward those who help them and punish those who hurt them (Rabin, 2001). Thus, responses are not based solely on individuals' interest in equity or fairness, but on their desire to rectify wrongs and reward good behavior. Punishment, in this view, is a form of reciprocity which may have evolved to stabilize cooperation (Fehr and Gächter, 2002; Boyd *et al.*, 2003).

Although in many cases predictions are similar, reciprocity and inequity-aversion models can be compared. In computerized situations, in which play against a computer (which lacks intentions) is compared with play against another person, people playing a computer are much less likely to reject low ultimatum game offers than in situations in which they play another human (Blount, 1995). This indicates an element of reciprocity in their actions. However, people will also take money away from others when income distributions are chosen at random (i.e., no intention can be attributed), indicating that equity concerns can outweigh reciprocity (or punishment) concerns (Dawes *et al.*, 2007). Thus, both inequity aversion and reciprocity appear to play a role in human decision making.

There is an enormous literature on these games, indicating that many other factors are at work. For instance, there is quite a bit of variation in how different cultures play the ultimatum game, although a broad survey of non-industrial cultures shows that no cultures play as if they were the *Homo economicus* parody, interested only in maximizing their own gains without paying attention to others' gains (Henrich *et al.*, 2001). Even in the culture in which proposers offered the least, the Machiguenga (a traditional, non-market society in the Amazonian rainforest), offers were well above the minimum. Individual variation plays a role as well. For instance, individuals who are in the lowest third of the population for self-esteem rarely respond to procedural inequity, indicating that personality or social factors affect individual responses (Wiesenfeld *et al.*, 2008). A recent study indicates that responder behavior in the ultimatum game is highly heritable (Wallace *et al.*, 2007), evidence which supports the presence of a genetic mechanism through which these reactions can be passed on. It is likely that personality, genes, and culture interact to create the range of responses that any one individual is likely to display.

Framing effects are important as well. For instance, offers drop – and these lower offers are accepted – if the role of the proposer is earned (in this case, through a high score on a quiz) rather than assigned randomly by the experimenter (Hoffman *et al.*, 1994). Social punishment drives offers in the other direction; proposers who know that their responders will be writing them a note after the division tend to offer higher amounts (Xiao and Houser, 2005). Interestingly, the responders actually accept lower amounts in these cases, but send back particularly vitriolic messages.

Although discussions of economic decision making tend to invoke images of conscious cognition, an emotional component may be involved. Robert Frank argues that showing a strong negative response to inequity, perhaps especially if it is irrational (in the sense of costing more than the benefit gained from rectifying the inequity), may be a powerful commitment device which increases the fairness of one's treatment. Such a negative response may carry a short-term cost, but this is more than recouped by the long-term gains of being treated fairly in the future (Frank, 1988). Recent experimental evidence supports this view. Players in the impunity game (an ultimatum game in which responders can only reject their allotment; responders cannot affect the proposer's outcome; see Box 19.2) do refuse their rewards, even though this increases both absolute and relative inequity (Yamagishi, 2007). Perhaps a part of people's reactions is to show others, and themselves, that they will not accept such treatment.

Recently, brain-imaging has been added to the arsenal of study tools for understanding humans' behavior in economic decision-making games. These studies corroborate this idea that emotional involvement is critical. Functional MRI, a technique for measuring the amount of oxygen in a given part of the brain (which correlates with activity) in real time, allows researchers to investigate what parts of the brain are activated during different decision-making tasks. These scans indicate activity in both conscious processing centers and areas of the brain correlated with emotion and impulse control during many of these decision making games (see, for example, Sanfey *et al.*, 2003).

For instance, in a prisoner's dilemma game (see Box 19.2), during unfair responses activation is seen in both the anterior insula, associated with emotional processing, and the dorsolateral prefrontal cortex (DLPFC), associated more with cognition (Sanfey *et al.*, 2003). Increased activation of the anterior insula is seen when individuals reject unfair offers, perhaps indicating a higher degree of emotional involvement. Correlating with earlier behavioral data (Blount, 1995), brain activation is different for people playing the game against another human as against a computer.

A more recent study found that interrupting the right DLPFC (but not the left) with repetitive transcranial magnetic stimulation (TMS) increases the rate of acceptance of low offers. This suggests that the DLPFC is critical in over-riding the self-interested impulse to accept all positive offers (Knoch *et al.*, 2006). Interestingly, in this case, behavior was the same in a computerized version of the task as in the human partner condition, indicating that the refusal of low offers may be based on reciprocity rather than inequity aversion.

WHY NON-HUMAN PRIMATES ARE RELEVANT TO INEQUITY

Humans are primates, sharing a taxonomic Order (Order *Primates*) with all other primates and a Family (Family *Hominidae*) with the other great apes (see Figure 19.1). Of these, we are most closely related to chimpanzees and bonobos, both members of our sister Genus, *Pan*. Generally, primates are divided in to Prosimians, which are the most phylogenetically ancient of the primates; the New World monkeys (or Platyrhines), which diverged from the line that led to humans approximately 40 million years ago (and include the capuchin monkey, which is discussed further below); the Old World monkeys (or Cattarhines), which diverged approximately 20 million years ago; and the great apes, of which humans are a member, along with chimpanzees, bonobos, gorillas, and orangutans.

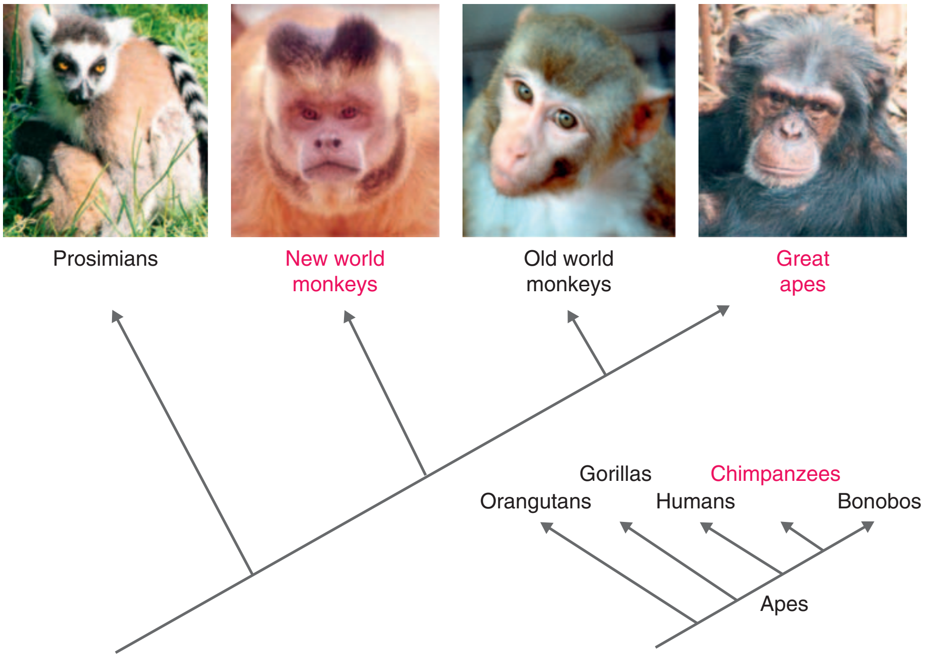


FIGURE 19.1 A taxonomic tree of the primate lineage, with species used in these studies indicated in red. Pictures indicate representative species of the four major taxonomic divisions of the Order Primates: Prosimians (a ring-tailed lemur female and her offspring), New World monkeys (an adult male capuchin monkey, one of the individuals used in these studies), Old World monkeys (an adult female rhesus macaque), and the Great Apes (an adult male chimpanzee, another individual used in these studies). An inset shows the relationships between the five members of the family Hominidae (the great apes). Photograph of the rhesus macaque by F. de Waal; all others by the author.

Studying non-human primates, our closest evolutionary relatives, tells us a great deal about how we behave. A technique known as *behavioral phylogeny* allows comparisons to be made between different species to extrapolate the likelihood of common descent for any behavior (Boehm, 1999). Species with different degrees of relatedness (such as an Old World monkey species and a New World monkey species) are evaluated for a behavior, following which an attempt is made to reconstruct when the behavior arose and how it developed over evolutionary time. Studying species which are more closely related yields more information – thus the utility of studying primates, and perhaps particularly great apes, when concerned with human behaviors.

Moreover, studying primates offers some advantages over studying humans. Within a group of primates, we typically know their full social history and, particularly in captivity, have more knowledge of (and control over) their lifetime experiences. Non-human species are also unencumbered with complex cultural institutions, such as churches, schools, and governments, so their responses are likely to indicate what primates have evolved to do (although primates do have impressive social learning abilities which undoubtedly affect their behavior; de Waal, 2001). By studying non-human primates' behavior, we learn the basis of human behavior and can begin to tease apart what is evolved and what is the result of our extensive culture. Finally, non-human primate social systems are simpler than human systems, if only because the number of other possible interactors is limited. This means that it is more likely that we can begin to understand all of the factors affecting their behavior.

The species discussed in this chapter are the chimpanzee, *Pan troglodytes*, and the brown, or tufted, capuchin monkey, *Cebus apella*. These species are ideal for comparison with humans in social and cognitive behavior. Both species have much larger brains and neocortices for their body size than is expected, based on the average brain-to-body ratio for both mammals and primates (Rilling and Insel, 1999). This implies a premium on intelligence, which, in social animals, is often assumed to arise due to the necessity of navigating a complex social environment in group-living species (Byrne and Whiten, 1988). Moreover, both of these species share several behaviors – food-sharing, reciprocity, and cooperation – that are somewhat unusual in primates, yet common in humans. Importantly, these behaviors may also be relevant to recognition of and response to inequity. Observations of natural behaviors in chimpanzees led to the early conclusion that a sense of justice might exist in their society (de Waal, 1991). For instance, chimpanzees will assist others who are victims of attempted theft (Figure 19.2).

FIGURE 19.2 Emily recruits support from the alpha male against the beta male, who has just attempted to take her food. Behaviors such as these indicate that chimpanzees may expect a certain level of equity in their daily lives. Photograph by the author.

FIGURE 19.3 A juvenile male eats from food in the possession of the alpha male while an adult female chimpanzee watches. Such close proximity may result in the female receiving food as well. Photograph by the author.

Food-sharing is rare outside of the mother/infant context in primates (Feistner and McGrew, 1989); however, both capuchins and chimpanzees commonly share food among adults and unrelated individuals (see Figure 19.3 for an example of begging and food-sharing among unrelated chimpanzees). In experiments with captive populations, capuchin monkeys not only food-share (de Waal, 1997a), but also tend to do so reciprocally (de Waal, 2000), more often sharing with those who shared with them. Food-sharing may also be a repayment for a service, as capuchins are more likely to share food with a partner who recently helped them in a cooperative task than with a partner who was inessential (de Waal and Berger, 2000). Finally, wild capuchin monkeys share food after successful hunts (Perry and Rose, 1994).

Chimpanzees also share food in reciprocal contexts (de Waal, 1989). For instance, chimpanzees are much more likely to share food with individuals who groomed them earlier in the day, and this tendency is much stronger for social partners who do not typically groom (de Waal, 1997b). This implies a scarcity effect, with more repayment due to an unusual behavior. Moreover, reciprocity works both ways; chimpanzees are also more likely to refuse to share with an individual who has not groomed them.

Chimpanzees also share food in the wild, typically sharing meat after a monkey hunt. This sharing may "buy" the sharer support, mating opportunities, or other social services from the recipient (Nishida, 1992; Rose, 1997; Mitani and Watts, 2001; but see Gilby, 2006 for an alternate view). One recent study indicates that male chimpanzees may also share difficult-to-obtain cultivated fruits with females, perhaps with future mating opportunities in mind (Hockings *et al*., 2007). These two contexts indicate that food-sharing may be more valuable when the food is a costly commodity than when it is one that can be obtained easily and individually. Outside of the food context, chimpanzees also share services (grooming, support) reciprocally (Mitani, 2006; Duffy *et al*., 2007).

Cooperation, closely linked to reciprocity and food-sharing, is also common in both species. For instance, a series of studies has been done on capuchin monkeys in which they must work together to pull in a heavily counterweighted tray in order to obtain food rewards. Capuchins are very good at intuitive cooperative tasks (Brosnan and de Waal, 2002), and seem to take their partner's behavior into account. Capuchins' performance is much lower when they cannot see their partner (Mendres and de Waal, 2000). They also appear to know that their partner is required to complete the task; monkeys paired with a partner who can leave the testing set-up at will pull predominantly when their partner is present in the testing area, and thus is also available to pull (Mendres and de Waal, 2000). Finally, capuchins are very sensitive to the distribution of rewards in a cooperative task. If rewards are sometimes clumped (monopolizable) and sometimes dispersed (spread apart, preventing monopolization), from the first exposure to the task, monkeys pull much less for clumped rewards (de Waal and Davis, 2002). This sensitivity to payoff indicates that they are aware of their partner both as a cooperator for pulling in the tray and as a competitor in dividing the bounty. Chimpanzees, too, successfully utilize a bar-pull apparatus to obtain food rewards (Crawford, 1937). As with capuchins, chimpanzees are sensitive to the social context, cooperating more frequently with socially tolerant partners (Melis *et al*., 2006a) and, when given the choice, recruiting tolerant partners (Melis *et al*., 2006b).

In the wild, both capuchin monkeys and chimpanzees hunt vertebrate prey, which requires a cooperative effort (Rose, 1997). Moreover, some groups of chimpanzees show remarkable coordination of hunting behavior, typically taking the same roles in the hunt over time (Boesch, 1994). Capuchins also cooperatively defend themselves against predators (Perry *et al*., 2003).

In all of these cases of cooperation, reciprocity, and food-sharing, individuals tend to share or cooperate based on what the other contributed. This indicates that the primates are concerned not only with their payoffs, but also with the effort involved in the other's action. Perhaps for any cooperative species, an accurate assessment of both effort (costs) and payoff (benefits) is required to ascertain when to cooperate (van Wolkenten *et al*., 2007). Thus, these other cooperative primate species may be very good models for investigating the evolutionary origins of the response to inequity and fairness in humans.

REACTIONS TO INEQUITY IN NON-HUMAN PRIMATES

The Evolution of Inequity Aversion

It is unlikely that a behavior as complex as inequity aversion arose *de novo*. Instead, it probably evolved through a series of stages that were each beneficial to performers in their own right, ultimately resulting in the sort of complex understanding of inequity and fairness that we see in humans (Brosnan, 2006a, 2006b). Thus, while other species may or may not show behaviors that are identical to humans', these behaviors are likely to be steps in the same evolutionary process. Obviously, we expect to see more similarities in species that are more closely related to humans (such as chimpanzees) than more distantly related ones.

Previously, I have proposed four distinct steps in the evolution of the inequity response as we see it in humans (Brosnan and de Waal, 2004a; Brosnan, 2006b). The first is an ability to notice when rewards differ between individuals. While this seems obvious from a human perspective, it is not at all certain in other species. This requires individuals to pay attention to another outside of a context in which that other individual is directly affecting them (i.e., not during a fight, a sexual encounter, etc). Moreover, it requires fairly advanced cognition to compare one's own rewards with those of another. However, it is likely that this ability arose in a different context than inequity. Many species socially learn, meaning that they acquire information or techniques from watching other individuals' actions or outcomes. This is an efficient way of acquiring information that changes frequently. Any species that learns socially, including chimpanzees and capuchin monkeys (Whiten, 1998; Brosnan and de Waal, 2004b), must pay attention to other individuals' rewards. Thus, they possess the capability to compare their rewards to those of others.

A second step is to react negatively to discrepancies in the rewards between one's self and another. Essentially, this requires individuals not only to notice the outcomes of another (the first step), but also to react negatively to this difference. In fact, this alone may provide a fitness benefit, if this negative response to inequity causes individuals to seek new partners who might be more equitable. For instance, both capuchins and chimpanzees hunt cooperatively, then share the prey after the kill (Rose, 1997). If one individual in a partnership consistently receives little or no meat after the hunt, hunting with a new partner who shares more generously would increase the first individual's meat consumption (and evolutionary fitness, or reproductive potential). Individuals need not intentionally seek out better sharers, nor do they need to understand their motives for switching partners, for this to enhance the actor's evolutionary fitness. The positive outcome will yield a fitness benefit to any individual who is inclined to "shop around" for a better partner.

The third step is to take deliberate action to rectify inequity towards the self, which humans do (see "Inequity aversion" above). This is the most complex of the levels of inequity. It requires an understanding of how one's actions will alter the outcome of the partner, and some ability to inhibit, to give up the immediate reward. Non-human primates can inhibit their behavior, even in the face of an immediate reward (Beran, 2002; Evans and Westergaard, 2006), and show basic understanding of how their actions will be perceived by others (Hare *et al.*, 2000).

The fourth and final step in the evolution of inequity responses is a response against overcompensation (Walster *et al.*, 1978), or advantageous inequity aversion (Fehr and Schmidt, 1999). In this case, individuals are uncomfortable when another individual receives less than they themselves do. This can be true both when the advantaged party is the self and when it is a third party (Andreoni *et al.*, 2002). Rectification usually takes the form of psychological leveling mechanisms (e.g. derogation, justification for receipt of the superior reward) or, more practically, material compensation.

Note that all four of these stages need to be developed in order to consider the response a sense of "fairness" or justice in the way that we mean it in humans (Brosnan, 2006a, 2006b). Particularly if a reaction to overcompensation is missing, the response is a one-sided response that is focused purely on the self. Generally, among humans there is an appeal to an "objective" level of fairness or justice by which actions are to be judged. (Of course, objectivity is still in the eye of the beholder.) Since it is difficult or impossible to determine a non-human's ideals, in these studies we examine purely behavior, not motivation, with the goal of determining situations in which the primates show evidence of one of these stages and, ultimately, a potential evolutionary trajectory from comparisons between these species.

Reactions to Disadvantageous Inequity

Eighty years ago, a comparative psychologist named Otto Tinklepaugh examined macaques' responses to violation of expectation (Tinklepaugh, 1928). Monkeys were shown a food, which was then hidden. Later, the researcher uncovered it and gave it to the monkey. Sometimes, he switched the food with a less desirable one without the monkey's knowledge. In these cases, the monkeys responded by becoming upset and refusing the less valuable foods. This demonstrated that these monkeys had expectations about what they should have received.

What Tinklepaugh did not do was investigate how they responded to a partner getting a better reward. Would the monkeys then form expectations based on what their partner received, and respond negatively if they got a less good reward than their partner? Our group has now performed this experiment. Our studies have investigated whether primates would be willing to complete an exchange interaction (return a token and accept a food reward) if their partners got a better food for the same "price," an exchange (see Figure 19.4 for a sketch of the test set-up). We tested capuchin monkeys and chimpanzees in pairs, with members of their social groups as their partners. In each case, one member of the pair, the partner, received a good reward (in this case a grape)

FIGURE 19.4 An adult female capuchin monkey completes an exchange with the experimenter, returning a token for which she will receive a food reward. An adult male from her social group looks on. Sketch by G. Bragg and F. de Waal (after a video still).

for completing an exchange with the experimenter, while the subject received a reward which was less valuable, but nonetheless typically desired (a piece of cucumber), for the same work – the exchange. These results were compared to a control in which both individuals exchanged to receive a cucumber, as well as several other conditions discussed below. We measured subjects' willingness to complete the exchange interaction and their willingness to consume the cucumber once they had received it. It turned out that subjects responded in each way about half of the time, so for the following analyses, unless otherwise noted, these categories are pooled and willingness to complete the interaction (return the token AND consume the food) is reported. In both species, subjects completed virtually all exchanges when both individuals got cucumber (the equitable situation), but failed to complete the interaction in as many as 80% of trials in which their partner got the (better) grape (Figure 19.5; Brosnan and de Waal, 2003; Brosnan, Schiff, and de Waal, 2005). Thus, both the ape and the monkey appear to be responding negatively to receiving less.

However, a few possibilities must be eliminated before this can be considered a reaction to inequity. For instance, it may just be "greed," or a desire for a better food that is available, rather than "envy," or a desire for a better food that another individual currently possesses (Brosnan and de Waal, 2004a; Wynne, 2004). To take this in to account we did an additional control, in which subjects were shown a grape, following which they exchanged and received only a cucumber piece. With the capuchins, this was done in two different ways. In the first, the subject was isolated, rather than being tested as a member of a pair, and grapes were placed in the partner's side of the testing chamber prior to each exchange. In this way, subjects saw the grapes pile up while they continued to receive only cucumber (Brosnan and de Waal, 2003). Subjects discriminated between this and the situation in which a partner received the grape, becoming more likely to accept the cucumber over time when no other monkey received a grape (Figure 19.6a). However, in this situation subjects were tested alone, while in all other conditions subjects were tested with a social partner. Although it seems minor, for a gregarious species such a change in the social environment may have large effects. Thus, we used a second method with the capuchins, as well as with the chimpanzees, in which subjects remained with their partner and each was shown a grape immediately before their exchange. Again, individuals were exposed to grapes in this testing situation, but no other subject received one. As when subjects were tested alone, subjects in both species discriminated between this and the situation in which a partner received the grape, becoming more likely to accept the cucumber over time when no other individual received a grape (see chimpanzee data and capuchin data in Figures 19.6b and 19.6c, respectively; Brosnan *et al.*, 2005; van Wolkenten *et al.*, 2007). This indicates that the partner receiving a grape is much more important than the presence of grapes alone.

A second possibility is that the subjects were responding negatively to the contrast between what they got in the current session and what they had received in previous sessions – a phenomenon sometimes called the frustration effect. Perhaps a monkey receiving a cucumber, but remembering the receipt of a grape in the last session, is responding negatively to the contrast with their previous experience, rather than the contrast with their partner's current experience. A study finding evidence for the frustration effect under these conditions, however, left out both the exchange task and a comparison between frustration and inequity, making it unclear what the critical

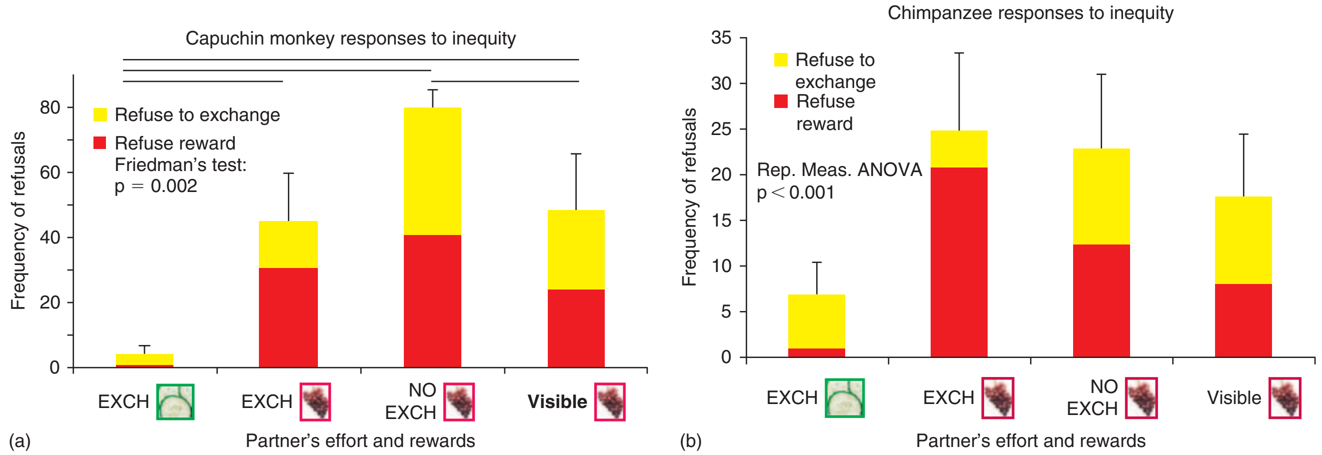


FIGURE 19.5 The distribution of refusals in the original conditions. The subject always exchanged for a cucumber, while the partner's rewards and effort required varied (indicated on the x-axis; cucumber or grape for reward, exchange or no exchange for effort). In the "Visible grape" condition, grapes were visible but no primate received one. Paler bars indicate refusal to return the token; darker bars indicate refusal to accept the reward. Standard error bars are for the total measure (refusal to return the token and refusal to accept the reward combined). Subjects were much more likely to refuse to participate when the partner received a better reward, either for the same or for less effort. Significant differences from *post hoc* tests are indicated by horizontal lines over bars. (a) Data for capuchin monkeys. Reprinted from Brosnan and de Waal (2003). (b) Data for chimpanzees. Reprinted from Brosnan *et al*. (2005).

factor is in the monkeys' responses (Brosnan and de Waal, 2006; Roma *et al*., 2006).

We recently completed a much larger study, replicating our original work with more conditions (discussed below) and more subjects. This allowed us to investigate how subjects' rewards in previous sessions affected their current responses. If frustration is a factor, then subjects should be more inclined to refuse to participate in sessions immediately following those in which they received grapes. However, we found no evidence that previous experience had any effect on their response; subjects were no more likely to refuse to participate if they had received a grape (frustration) versus a cucumber (no frustration) on the previous test. Thus, if frustration does play a role, it is swamped by the response to inequity (Figure 19.7; van Wolkenten *et al*., 2007).

Context is also critical in responses to inequity. Several studies have looked for inequity using a task in which foods are simply handed to the subjects, without any sort of task. One study found an increase in the rate of food consumption when the subject got a less desirable food than their partner received (Bräuer *et al*., 2006), but in none of these studies did subjects respond to inequity by refusing to participate (Dindo and de Waal, 2006; Dubreuil *et al*., 2006). However, one of these studies utilized the same subjects who responded to inequity in three other studies involving a task (both of the exchange tasks discussed above and a study discussed below in "Reactions to an 'unfair' partner"). Thus, the lack of response to inequity in situations without a task is not a general lack of response to inequity, but instead appears to be specific to situations involving a specific task or effort (Dindo and de Waal, 2006).

If responses to inequity arose in the context of cooperative effort, responses to inequity would not be expected in situations with no required effort. When effort is involved, payoff equity becomes much more critical. To place this in an ecological context, a monkey who learns to forage in a certain tree after seeing another individual obtain better food there (the equivalent of being handed a reward) is exhibiting social learning, not a response to inequity. The fact that one found a better cache of food should not alter their relationship. However, if the same two monkeys hunt cooperatively, and one individual takes a much larger portion of the ensuing kill, then the disadvantaged partner does need to evaluate their relationship and possibly find a new hunting partner who shares the prey (payoff) more equitably. Joint efforts must be rewarded jointly, or the cooperative effort will fall apart.

Of course, if effort is required, it is possible that differences in effort may generate differences in response. In the second exchange study (van Wolkenten *et al*., 2007), we included conditions in which subjects had to exchange only a single time and others in which subjects had to exchange three times ("high effort" condition) to get the reward. Capuchin monkeys are sensitive to effort, responding less frequently and more slowly when they must expend more effort to obtain a reward. This is also modulated by reward value. When working for good rewards (grapes), monkeys are somewhat indifferent to the amount of effort required,

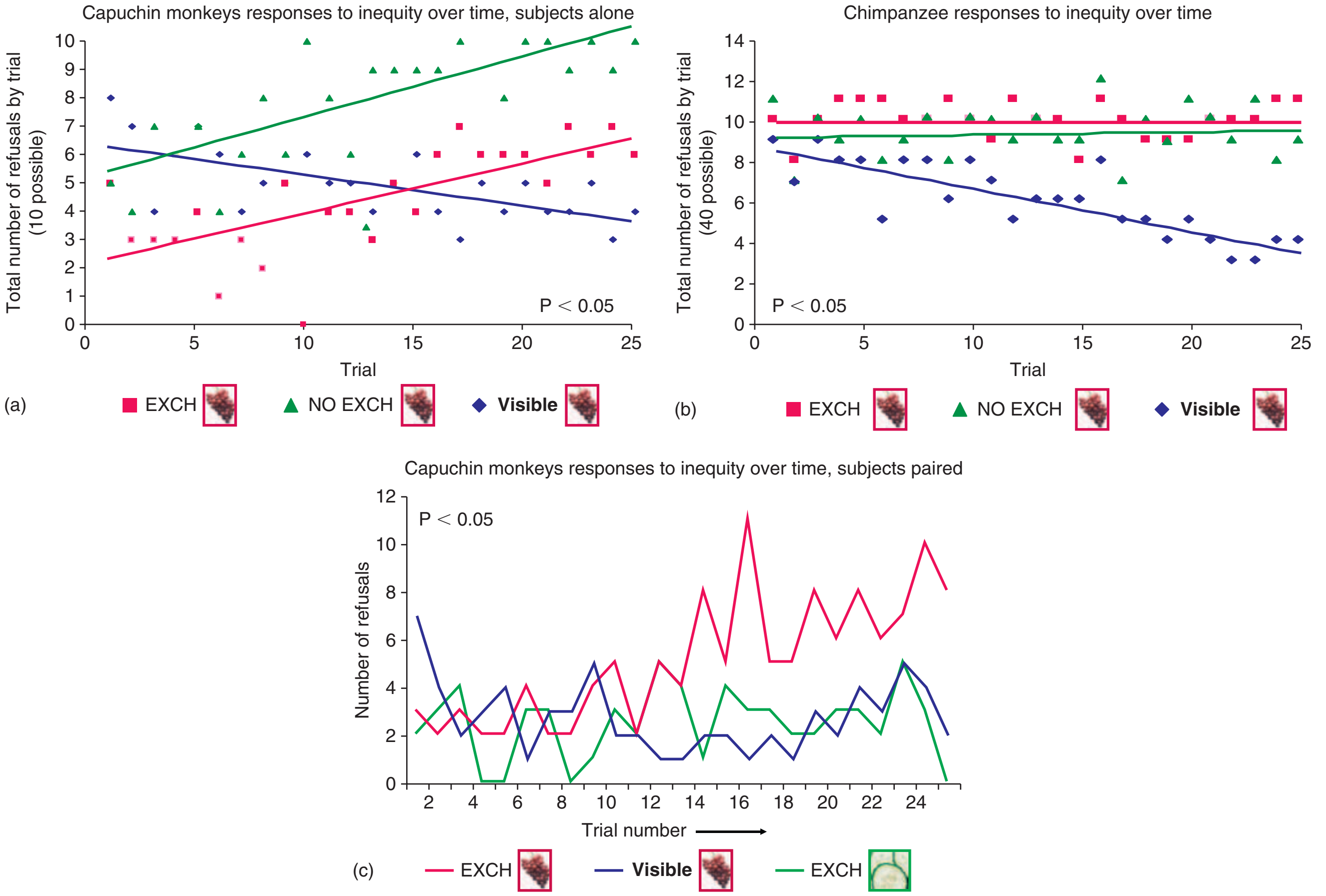


FIGURE 19.6 The total number of refusals to exchange by trial in the three test conditions (explained in Figure 19.5). Subjects' willingness to complete the interactions changed over time dependent upon whether grapes were only visible (Visible grape condition) or were given to a partner (Exchange grape, No exchange grape). Figure 19.6c lacks the No exchange grape condition, but includes the control equity condition in which both individuals receive cucumbers for comparison. This indicates that the refusals depicted in Figure 19.4 are due not to the mere presence of grapes, but to the fact that another individual received them. (a) Data for capuchin monkeys tested alone (in grape visible condition). Reprinted from Brosnan (2006b). (b) Data for chimpanzees tested in pairs (all conditions). Reprinted from Brosnan *et al*. (2005). (c) Data for capuchin monkeys tested in pairs (all conditions). Reprinted from van Wolkenten *et al*. (2007).

completing the task equally often in situations requiring both low and high effort. However, when reward value is low (cucumbers), monkeys are very sensitive to effort, completing many fewer trials when the effort required is high (van Wolkenten *et al*., 2007). Note that this means that capuchins are only willing to forego a reward if it is not particularly valuable. Thus, capuchins are sensitive to the price of "disdaining" a reward.

However, the above simply means that the monkeys are sensitive to their own level of effort, and not necessarily to the effort of their partner in comparison to their own. While earlier studies indicated the importance of effort (Brosnan and de Waal, 2003), later studies have cast doubt on this (Fontenot *et al*., 2007). Thus, we also compared the situation in which the both individuals had to work (exchange) for the same foods, but the effort – or number of exchanges – required differed. In this case, subjects did not refuse more frequently when their partner worked less hard (exchanged only once), indicating that variations based on effort may be due to one's own assessment of effort, rather than comparisons with the partner's effort. However, this involved only small differences in effort (one exchange versus three), so potentially larger differences in effort, or a different context, could elicit a different response. Given that inequity may have arisen in a cooperative context, a cooperative task may be the most appropriate context for continued investigation (see "Reactions to an 'unfair' partner" below for such a task).

One final factor that may play a role is the relationship between the individual in question and his or her social partner. In the case of capuchin monkeys, the identity of the social partner did not appear to affect their responses. However, chimpanzees from a long-term stable social group were much less likely

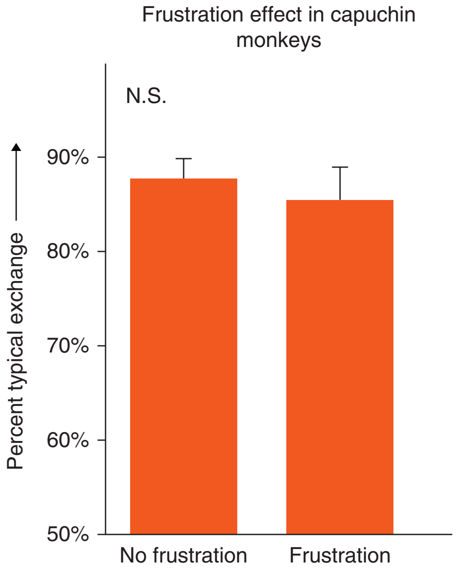


FIGURE 19.7 The frequency of subjects' refusals to exchange for cucumber dependent upon whether the subject received grape (frustration) or cucumber (no frustration) in the previous session. Subjects were no more likely to refuse to exchange in the frustration condition, indicating that their response is not simply a contrast between the current cucumber and previously available rewards. Reprinted from van Wolkenten *et al*. (2007).

to respond negatively when their partner, from that social group, received a better reward, than were those from a group which had been formed much more recently (Figure 19.8; Brosnan *et al*., 2005). Although time together and the strength of their relationships are confounded here, this is critical evidence that some aspect of social environment affects the inequity response. This is true in humans as well. Humans in close, communal relationships show much less response to inequity than do those in more distant, or more contingent, relationships (Clark and Grote, 2003). Perhaps the chimpanzees which had lived together literally since birth simply had better relationships than those introduced as adults, and do not respond to inequity. In fact, one other study, which found that chimpanzee responders in an analog to a limited form ultimatum game typically accepted all rewards, also used subjects from a stable social group (Jensen *et al*., 2007). The social context of inequity is a critical element, and one which is currently being examined in more detail.

In the economic understanding of inequity aversion, a negative reaction to inequity should alter payoffs in a way that maximizes equity (Henrich, 2004). Thus, a key difference between how the primates responded and the predictions of the inequity aversion models is that refusing the cucumber in this situation is economically irrational. If subjects are inequity averse, they should never react in a way that increases both absolute and relative inequity (this is irrational), yet this is exactly what the primates are doing. Subjects did not change their immediate outcome, their long-term outcome, or their partner's outcome when they refused to accept the lesser reward. Thus, their behavior is difficult to explain using existing models. Of course, even humans may respond in this "irrational" way, as, for instance, in the impunity game (see Box 19.2 and Yamagishi, 2007; however, see also Bolton and Zwick, 1995, for the opposite finding).

Why might people and primates respond in this way? This response cannot be explained by either inequity aversion or reciprocity. First, it is possible that in the primates' normal interactions, their responses *would* alter payoffs. In the experiment, subjects were trapped for 25 trials, whereas in the wild, subjects could cease interaction with an inequitable partner and go to find a more equitable one. Second, in a cooperative task (in which individuals work together for joint rewards), if one individual stops cooperating, neither individual can achieve the payoff, essentially rendering the payoff equal (that is, nothing; see "Reactions to an 'unfair' partner," below). Finally, this response may be, in both humans and non-human primates, an emotional response that functions as a commitment device (Yamagishi, 2007). Subjects may respond "in spite of themselves" in the experimental task, even though there is nothing to be gained in this specific context. The hypothesis that this response is driven by emotion in non-humans as well as humans is one that deserves much more investigation.

Reactions to Advantageous Inequity

A critical component of "fairness" includes responding to situations which are disadvantageous to another individual. In the preceding studies, individuals disliked being treated inequitably as compared to their partners, but none of the subjects responded differently if they were the advantaged partner. Capuchin monkeys frequently allowed their partner to take their cucumber pieces (passive sharing), but no monkey ever allowed a partner to take his or her grape (Brosnan and de Waal, 2003). Among chimpanzees, in less than 1% of cases (5 of 2000) did an individual who received grape allow the cucumber-receiving partner to have the grape (Brosnan *et al*., 2005). This rate of sharing is much lower has been seen by these same individuals in other studies (de Waal, 1997b).

However, these studies are not ideal for investigating how individuals respond to advantageous inequity. The food pieces are small and (by design)

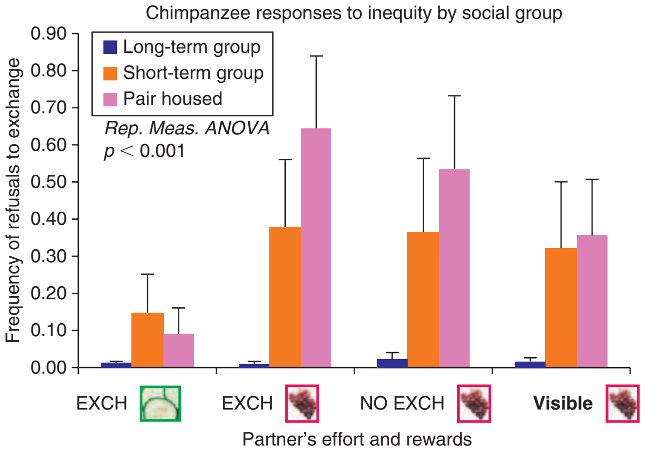


FIGURE 19.8 This figure further breaks down the data presented in Figure 19.5b, indicating that responses varied greatly dependent upon the social group. Chimpanzees from a long-term social group were much less likely to respond to inequity than those from shorter-term or pair-house situations, indicating that subjects' relationships or other social factors influence the response to inequity. Reprinted from Brosnan *et al.* (2005).

difficult to share. Thus, rectifying inequity requires the advantaged individual to give up the entire grape – essentially rendering them disadvantaged. We recently examined the willingness of chimpanzees to give food to a partner (see Chapter 18 for more details on methods and results). Subjects could choose from two options, one of which brought food to them and one of which which brought food to both them and a partner (prosocial option). Rewards were identical, and subjects received the same reward no matter what they chose. We compared subjects' responses when paired with another individual to their responses when alone, to avoid the possibility that subjects chose the "prosocial," or two-food, option due to the presence of additional food.

In fact, subjects were indifferent to their partner's outcomes. None of the 18 subjects from two different facilities ever showed a significant tendency to reward partners more when there was a partner present than when next to an empty cage; nor was there an effect overall (Silk *et al.*, 2005). Another similar study with a third group of chimpanzees found the same results (Jensen *et al.*, 2006). However, it is possible that the subjects were simply so excited by the prospect of receiving food that they made a choice for themselves without attending to the option of rewarding a partner.

Thus, we completed a separate study (the study order was counterbalanced between facilities) with the same subjects in which chimpanzees could make *two* choices, in any order, one of which rewarded just them and one of which rewarded their partner. In this case, while the majority of subjects were again indifferent, one female chose the prosocial option at a significantly higher rate when next to her partner than when next to an empty cage, and three other subjects from the same facility showed a tendency in this direction (out of 18 subjects). Moreover, on their initial session, the subjects in this facility chose the prosocial option more when a partner was present than absent (Vonk *et al.*, 2008; see also Chapter 18 of this volume).

Thus, while in these tests chimpanzees have been seen to be predominantly disinterested in their partner's rewards, there is a hint that they may have the capacity to make prosocial choices. In other situations, enculturated (i.e., human-reared) chimpanzees respond in a similarly helpful fashion to that of small children in a simple helping task (Warneken and Tomasello, 2006), and non-enculturated chimpanzees will assist both humans and other chimpanzees in other helping tasks (Warneken *et al.*, 2007). These helping behaviors, involving assistance rather than the donation of food, are more similar to the sorts of situations in which behaviors that appear empathetic have been reported previously (de Waal, 2006). Given the relative rarity of sharing of vegetable foods (as opposed to meat) in chimpanzees, it may not be surprising that behavior is different in experimental settings in which food is the reward versus those which involve only helping behavior. Social environment, too, may be critical, as cooperatively breeding marmosets do show prosocial behavior in food contexts (Burkart *et al.*, 2007).

It is likely that chimpanzees, at least, do show some prosocial behavior in some social contexts, although further research is needed to clarify what these contexts are (Brosnan, 2008; see also Chapter 18 of this volume). Whether this prosocial capacity is sufficiently well developed to manifest as a full sense of fairness in chimpanzees is still unknown. Certainly in the experimental exchange study, in which subjects responded negatively when they got less food than a partner, neither species showed altered behavior for the advantaged partner. Until further research is done, involving different contexts and different costs and benefits, it is best to avoid assuming that chimpanzees, or other primates, are concerned with overcompensation.

Reactions to an "Unfair" Partner

While previous work on distributional inequity focused on the distribution of rewards, more recent work on cooperation for unequal rewards shows that, in cooperative contexts, subjects may be more attuned to their partner's behavior than to the rewards themselves. Capuchins are very sensitive to both the social context and physical aspects of cooperation. In fact, on their first exposure to a mutualistic task in which individuals must work together to pull in a tray which brings both food, subordinate monkeys will refuse to cooperate for shared rewards if those rewards are monopolizable (de Waal and Davis, 2002).

In the previous exchange experiments, the monkeys' unequal rewards were allocated by the experimenter. The partner did not affect the outcome, and neither individual could do anything to affect the other's reward. This is likely not the case in natural situations; thus, in the current experiment, the subjects controlled the allocation of rewards. This was done using a bar-pull task, which has been used extensively to test cooperation and altruism in capuchin monkeys (see "Why non-human primates are relevant to inequity," above). A tray is weighted such that no capuchin can pull it in alone, but, working together, two individuals can bring it in. On the tray are two rewards, one for each monkey. Sometimes these rewards were the same (equal condition) and sometimes they were different (unequal condition). Subjects were not separated for this task, so they had to work out for themselves which monkey would pull for which reward, meaning that the monkeys themselves controlled the payoff distribution, or allocation of rewards. Note that if either individual refused to participate, the bar-pull tray could not be pulled in and neither got "paid."

In this situation, the distribution of the rewards did not affect a pair's cooperative success. However, the behavior of the partner strongly affected their responses. In the unequal condition (different rewards), some monkeys dominated the better reward, never giving their partners a chance to obtain it. Other monkeys were more egalitarian, taking the better reward only about half of the time in this unequal condition (Figure 19.9a). Pairs in which both individuals shared in the better rewards were almost three times as successful at the cooperative task as compared to those pairs in which one member consistently claimed the better reward (Figure 19.9b). Interestingly, this was true in both the unequal *and* the equal conditions. Monkeys whose partners took the better reward in the unequal condition quit cooperating in all conditions, whether or not there was a possibility of receiving a less good reward. An analysis of rewards obtained shows that, due to the much higher rate of cooperation, the egalitarian strategy resulted in far more rewards for both monkeys (Brosnan *et al.*, 2006).

It appears that these monkeys are reacting against unfair partners, rather than attempting to maximize reward intake. Monkeys simply did not cooperate with individuals who dominated the better rewards, even when the payoffs would have been the same. From a reward maximization perspective, this makes no sense, as subjects always got one of the rewards, so when rewards were equal, both individuals benefited equally. In fact, monkeys who failed to cooperate were no less likely to come over and observe the food rewards, so presumably they could have evaluated each trial independently. Perhaps they improve their long-term outcomes by avoiding individuals who typically do not behave equitably. This, too, hints at an emotional response against a partner.

Chimpanzees are likely to behave in a similar fashion, reacting against partners who do not treat them equitably. Although the same study has not been completed, work with cooperative tasks in chimpanzees shows that they are sensitive to partner behavior. Chimpanzees are much more likely to cooperate with individuals who tolerantly share food with them (Melis *et al.*, 2006a) and actively recruit tolerant partners (Melis *et al.*, 2006b) for cooperative enterprises. Future studies which further investigate monkeys' and apes' responses in cooperative tasks will clarify how sensitivity to context evolved, and whether it typically functions to increase payoffs over the long run.

Comparison of Human and Non-human Primate Responses to Inequity

Capuchin monkeys and chimpanzees both respond very negatively to receiving less benefit than a partner. This sort of reaction is similar to one we see in humans, as in the ultimatum game or the impunity game, in which people will turn down an absolute gain if the outcome is not equitable (Güth *et al.*, 1982), sometimes even when the response does not result in more equity between partners (Yamagishi, 2007). However, the reaction in non-human primates is not exactly the same as that seen in humans. The primary difference is that non-human primates do not respond to receiving more than their partner. Humans seem to treat "fairness" as a norm, while non-human primates appear more interested in their own outcomes. Of course, humans are not always interested in their fellow man's well-being, either, and do not always respond to overcompensation (see "Inequity aversion"). Future work with primates may uncover situations in which primates, too, are concerned about equity for others. Studying these situations would shed light on the contexts which elicit this response in humans.

Even with this difference, however, similarities predominate. In both humans and non-human primates, the social partner's actions appear to be more important than the actual payoff. Capuchin monkeys will accept an unequal payoff in the short term if the long-term payoffs are similar (equitable). This reaction

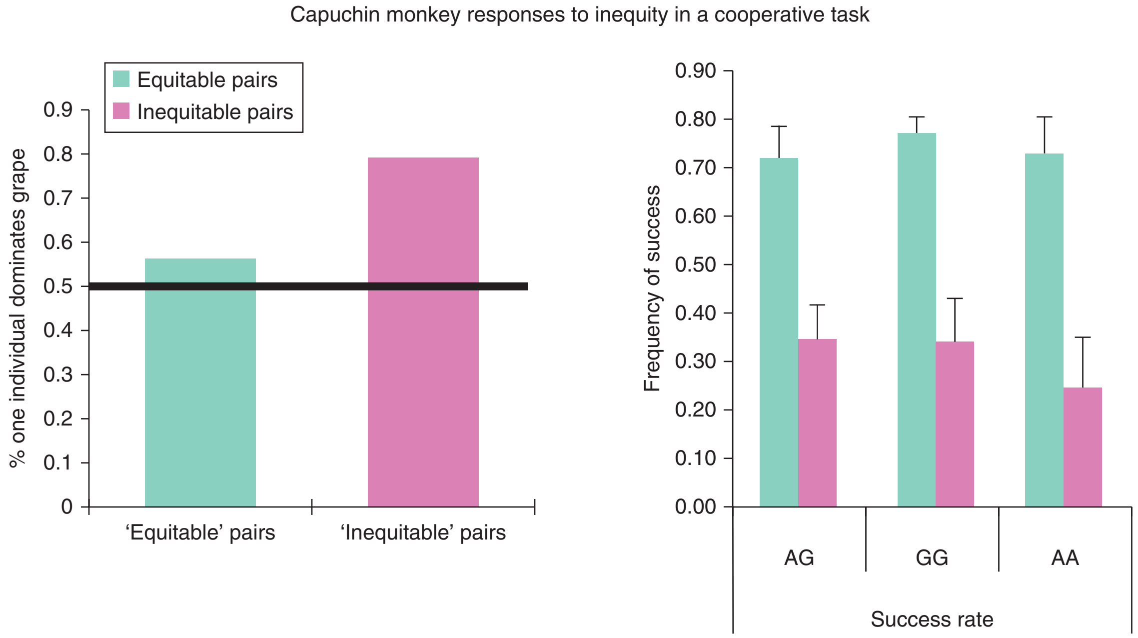


FIGURE 19.9 The frequency of successful cooperation for "equitable" and "inequitable" pairs in an unequally rewarded cooperation task. Some pairs (equitable) took turns receiving the higher-value reward, while in other pairs (inequitable) one individual claimed most of the better rewards. Equitable pairs were almost three times as likely to successfully cooperate in all conditions, including those with equitable rewards. AG = one subject gets apple, one subject gets grape; GG = both subjects get grape; AA = both subjects get apple. Reprinted from Brosnan *et al*. (2006).

promotes cooperation in situations in which immediate payoffs are not identical or one player makes a "mistake," behaving non-cooperatively when he or she either intended to cooperate or did not understand the task. Individuals who view the interaction through a more long-term lens will ultimately maintain more cooperative opportunities. Likewise, social context is undoubtedly important for both humans and non-human primates. Humans and chimpanzees both appear more likely to tolerate inequity in close relationship than in those which are more distant.

Non-human primates show several hallmarks of an emotional reaction to inequity. For instance, primates respond to distributional inequity in ways that lower their payoff, and respond negatively to a partner who has treated them inequitably, even in situations in which rewards are equal. Given what we know about the human response from imaging studies, it seems likely that these reactions are processed in emotion centers of the brain in non-human primates as well. Future studies that are able to investigate brain activity in awake primates interacting socially with other primates will help to clarify these questions.

Much work remains to be done before we fully understand the evolution of fairness. The experiments comparing monkeys, apes, and humans are not typically identical. This leaves open the possibility of secondary social or experimental factors influencing responses differently in the two species. For instance, non-human primate partners are always paired non-anonymously with a social partner from their group (they become too anxious to participate if paired with a stranger). In experimental economics games, on the other hand, people are typically tested with no cues as to their social partner's identity, or see them only on a screen. However, even the threat of anonymous, written, impersonal negative social feedback increases offers (Xiao and Houser, 2005), implying that results might be quite different in games which include social interaction.

CONCLUSIONS

Like humans, primates respond negatively to inequity directed against themselves. As with humans, these responses are difficult to explain with a single model of inequity aversion or reciprocity. Reciprocity may explain why, in the cooperative bar-pull study, monkeys were willing to accept a lesser reward on some trials if overall payoffs were the same. Inequity aversion may explain why a capuchin would refuse to cooperate when a partner routinely claims the better reward in an unequal payoff situation. Yet, refusing a reward in the exchange tasks actually increases relative inequity, which violates the inequity aversion models. This behavior, which occurs in humans as well as capuchin monkeys and chimpanzees, is difficult to explain. Perhaps for all three species, refusal to participate can be explained not only by inequity aversion and reciprocity, but also as a commitment device. Clearly, more research is needed to understand the mechanism for rejection in non-human primates, and to better understand the evolution of the behavior.

Non-human primates provide a valuable model for studying human behavior, including reactions to inequity. Responses to inequity, for instance, likely evolved in a common ancestor to humans, apes, and monkeys (and are possibly more ancient than this); thus, understanding how such a response manifests and the factors which affect it in other species helps us to understand how the reaction evolved into what we see in humans.

Acknowledgments

I thank the National Science Foundation (SES 0729244) and am grateful for an NIH IRACDA grant to Emory University for support.

References

Andreoni, J., Brown, P.M., and Vesterlund, L. (2002). What makes an allocation fair? Some experimental evidence. *Games Econ. Behav.* 40, 1–24.

Axelrod, R. (1984). *The Evolution of Cooperation*. New York, NY: Basic Books.

Beran, M.J. (2002). Maintenance of self-imposed delay of gratification by four chimpanzees (*Pan troglodytes*) and an orangutan (*Pongo pygmaeus*). *J. Gen. Psychol.* 129, 49–66.

Bernhard, H., Fischbacher, U., and Fehr, E. (2006). Parochial altruism in humans. *Nature* 442, 912–915.

Blount, S. (1995). When social outcomes aren't fair: the effect of causal attributions on preferences. *Org. Behav. Hum. Dec. Proc.* 63, 131–144.

Boehm, C. (1999). *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*. Cambridge, MA: Harvard University Press.

Boesch, C. (1994). Cooperative hunting in wild chimpanzees. *Animal Behav.* 48, 653–667.

Bolton, G.E. and Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *Am. Econ. Rev.* 90, 166–193.

Bolton, G.E. and Zwick, R. (1995). Anonymity versus punishment in Ultimatum Game bargaining. *Games Econ. Behav.* 10, 95–121.

Boyd, R., Gintis, H., Bowles, S., and Richerson, P.J. (2003). The evolution of altruisitic punishment. *PNAS* 100, 3531–3535.

Bräuer, J., Call, J., and Tomasello, M. (2006). Are apes really inequity averse? *Proc. R. Soc. Lond. B* 273, 3123–3128.

Brosnan, S.F. (2006a). At a crossroads of disciplines. *Social Justice Res.* 19, 218–227.

Brosnan, S.F. (2006b). Nonhuman species' reactions to inequity and their implications for fairness. *Social Justice Res.* 19, 153–185.

Brosnan, S.F. (2008). Inequity and prosocial behavior in chimpanzees. In: E.V. Lonsdorf, S. Ross, and T. Matsuzawa (eds), *The Mind of the Chimpanzee*. Chicago, IL: University of Chicago Press, (in press)..

Brosnan, S.F. and de Waal, F.B.M. (2002). A proximate perspective on reciprocal altruism. *Hum. Nature* 13, 129–152.

Brosnan, S.F. and de Waal, F.B.M. (2003). Monkeys reject unequal pay. *Nature* 425, 297–299.

Brosnan, S.F. and de Waal, F.B.M. (2004a). Reply to Henrich and Wynne. *Nature* 428, 140.

Brosnan, S.F. and de Waal, F.B.M. (2004b). Socially learned preferences for differentially rewarded tokens in the brown capuchin monkey, *Cebus apella*. *J. Comp. Psychol.* 118, 133–139.

Brosnan, S.F. and de Waal, F.B.M. (2006). Partial support from a non-replication: comment on Roma, Silberberg, Ruggiero, and Suomi. *J. Comp. Psychol.* 120, 74–75.

Brosnan, S.F., Schiff, H.C., and de Waal, F.B.M. (2005). Tolerance for inequity may increase with social closeness in chimpanzees. *Proc. R. Soc. Lond. B* 1560, 253–258.

Brosnan, S.F., Freeman, C., and de Waal, F.B.M. (2006). Partner's behavior, not reward distribution, determines success in an unequal cooperative task in capuchin monkeys. *Am. J. Primatol.* 68, 713–724.

Burkart, J., Fehr, E., Efferson, C., and van Schaik, C.P. (2007). Other-regarding preferences in a non-human primate: common marmosets provision food altruistically. *PNAS* 104, 19762–19766.

Byrne, R.W. and Whiten, A. (1988). *Machiavellian Intelligence*. Oxford: Clarendon Press.

Camerer, C. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Clark, M.S. and Grote, N.K. (2003). Close relationships. In: T. Millon and M.J. Lerner (eds), *Handbook of Psychology: Personality and Social Psychology*, Vol. 5. New York, NY: John Wiley & Sons, pp. 447–461.

Crawford, M. (1937). The cooperative solving of problems by young chimpanzees. *Comp. Psychol. Monogr.* 14, 1–88.

Dana, J.D., Weber, R.A., and Kuang, J. (2003). Exploiting moral wriggle room: behavior inconsistent with a preference for fair outcomes. Carnegie Mellon Behavioral Decision Research Working Paper No. 349, http://ssrn.com/abstract=400900.

Dawes, C.T., Fowler, J.H., Johnson, T. *et al.* (2007). Egalitarian motives in humans. *Nature* 446, 794–796.

de Waal, F.B.M. (1989). Food sharing and reciprocal obligations among chimpanzees. *J. Hum. Evol.* 18, 433–459.

de Waal, F.B.M. (1991). The chimpanzee's sense of social regularity and its relation to the human sense of justice. *Am. Behav. Scientist* 34, 335–349.

de Waal, F.B.M. (1997a). Food transfers through mesh in brown capuchins. *J. Comp. Psychol.* 111, 370–378.

de Waal, F.B.M. (1997b). The chimpanzee's service economy: food for grooming. *Evol. Hum. Behav.* 18, 375–386.

de Waal, F.B.M. (2000). Attitudinal reciprocity in food sharing among brown capuchin monkeys. *Animal Behav.* 60, 253–261.

de Waal, F.B.M. (2001). *The Ape and the Sushi Master*. New York, NY: Basic Books.

de Waal, F.B.M. (2006). *Primates and Philosophers*. Princeton, NJ: Princeton University Press.

de Waal, F.B.M. and Berger, M.L. (2000). Payment for labour in monkeys. *Nature* 404, 563.

de Waal, F.B.M. and Davis, J.M. (2002). Capuchin cognitive ecology: cooperation based on projected returns. *Neuropsychologia* 1492, 1–8.

Diekmann, K.A., Samuels, S.M., Ross, L., and Bazerman, M.H. (1997). Self-interest and fairness in problems of resource allocation: allocators versus recipients. *J. Pers. Social Psychol.* 72, 1061–1074.

Dindo, M. and de Waal, F.B.M. (2006). Partner effects on food consumption in brown capuchin monkeys. *Am. J. Primatol.* 69, 1–6.

Dubreuil, D., Gentile, M.S., and Visalberghi, E. (2006). Are capuchin monkeys (*Cebus apella*) inequity averse? *Proc. R. Soc. Lond. B* 273, 1223–1228.

Duffy, K.G., Wrangham, R.W., and Silk, J.B. (2007). Male chimpanzees exchange political support for mating opportunities. *Curr. Biol.* 17, R586.

Evans, T.A. and Westergaard, G.C. (2006). Self-control and tool use in tufted capuchin monkeys (*Cebus apella*). *J. Comp. Psychol.* 120, 163–166.

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature* 415, 137–140.

Fehr, E. and Rockenbach, B. (2003). Detrimental effects of sanctions on human altruism. *Nature* 422, 137–140.

Fehr, E. and Schmidt, K.M. (1999). A theory of fairness, competition, and cooperation. *Q. J. Economics* 114, 817–868.

Feistner, A.T.C. and McGrew, W.C. (1989). Food-sharing in primates: a critical review. In: P.K. Seth and S. Seth (eds), *Perspectives in Primate Biology*, Vol. 3. New Delhi: Today and Tomorrow's Printers and Publishers, pp. 21–36.

Fontenot, M.B., Watson, S.L., Roberts, K.A., and Miller, R.W. (2007). Effects of food preferences on token exchange and behavioural responses to inequality in tufted capuchin monkeys, *Cebus apella*. *Animal Behav.* 74, 487–496.

Frank, R.H. (1988). *Passions Within Reason: The Strategic Role of the Emotions*. New York, NY: W.W. Norton & Co.

Gilby, I.C. (2006). Meat sharing among the Gombe chimpanzees: harassment and reciprocal exchange. *Animal Behav.* 71, 953–963.

Güth, W., Schmittberger, R., and Schwartze, B. (1982). An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Org.* 3, 367–388.

Hare, B., Call, J., Agnetta, B., and Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behav.* 59, 771–785.

Henrich, J. (2004). Inequity aversion in capuchins. *Nature* 428, 139.

Henrich, J., Boyd, R., Bowles, S. *et al*. (2001). In search of *Homo Economicus*: behavioral experiments in 15 small-scale societies. *Am. Econ. Rev.* 91, 73–78.

Hockings, K.J., Humle, T., Anderson, J.R. *et al*. (2007). Chimpanzees share forbidden fruit. *PLoS ONE* 2, e886, doi:10.1371/journal.pone.0000886..

Hoffman, E., McCabe, K., Shachat, K., and Smith, V. (1994). Preferences, property rights and anonymity in bargaining games. *Games Econ. Behav.* 7, 346–380.

Jensen, K., Hare, B., Call, J., and Tomasello, M. (2006). What's in it for me? Self-regard precludes altruism and spite in chimpanzees. *Proc. R. Soc. Lond. B* 273, 1013–1021.

Jensen, K., Call, J., and Tomasello, M. (2007). Chimpanzees are rational maximizers in an Ultimatum Game. *Science* 318, 107–109.

Kahneman, D., Knetsch, J.L., and Thaler, R. (1986). Fairness as a constraint on profit seeking: entitlements in the market. *Am. Econ. Rev.* 76, 728–741.

Knoch, D., Pascual-Leone, A., Meyer, K. *et al*. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314, 829–832.

Loewenstein, G.F., Thompson, L., and Bazerman, M.H. (1989). Social utility and decision making in interpersonal contexts. *J. Pers. Social Psychol.* 57, 426–441.

Melis, A.P., Hare, B., and Tomasello, M. (2006a). Engineering cooperation in chimpanzees: tolerance constraints on cooperation. *Animal Behav.* 72, 275–286.

Melis, A.P., Hare, B., and Tomasello, M. (2006b). Chimpanzees recruit the best collaborators. *Science* 311, 1297–1300.

Mendres, K.A. and de Waal, F.B.M. (2000). Capuchins do cooperate: the advantage of an intuitive task. *Animal Behav.* 60, 523–529.

Mitani, J.C. (2006). Reciprocal exchange in chimpanzees and other primates. In: P. Kapeller and C.P. van Schaik (eds), *Cooperation in Primates and Humans: Evolution and Mechanisms*. Berlin: Springer, pp. 111–113.

Mitani, J.C. and Watts, D.P. (2001). Why do chimpanzees hunt and share meat? *Animal Behav.* 61, 915–924.

Nishida, T., Hasegawa, T., Hayaki, H. *et al*. (1992). Meat-sharing as a coalition strategy by an alpha male chimpanzee? In: T. Nishida, W.C. McGrew, P. Marler *et al*. (eds), *Topics in Primatology: Human Origins*, Vol. 1. Tokyo: University of Tokyo Press, pp. 159–174.

Perry, S. and Rose, L. (1994). Begging and transfer of coati meat by white-faced capuchin monkeys, Cebus capucinus. *Primates* 35, 409–415.

Perry, S., Manson, J.H., Dower, G., and Wikbert, E. (2003). White-faced capuchins cooperate to rescue a groupmate from a Boa constrictor. *Folia Primatologica* 74, 109–111.

Rabin, M. (2001). Incorporating fairness into game theory and economics. *Am. Econ. Rev.* 83, 1281–1301.

Rilling, J.K. and Insel, T. (1999). The primate noecortex in comparative perspective using magnetic resonance imaging. *J. Hum. Evol.* 16, 191–233.

Roma, P.G., Silberberg, A., Ruggiero, A.M., and Suomi, S.J. (2006). Capuchin monkeys, inequity aversion, and the frustration effect. *J. Comp. Psychol.* 120, 67–73.

Rose, L.M. (1997). Vertebrate predation and food-sharing in *Cebus* and *Pan*. *Intl J. Primatol.* 18, 727–765.

Sanfey, A.G., Rilling, J.K., Aronson, J.A. *et al*. (2003). The neural basis of economic decision-making in the Ultimatum game. *Science* 300, 1755–1758.

Silk, J.B., Brosnan, S.F., Vonk, J. *et al*. (2005). Chimpanzees are indifferent to the welfare of unrelated group members. *Nature* 437, 1357–1359.

Tinklepaugh, O.L. (1928). An experimental study of representative factors in monkeys. *J. Comp. Psychol.* 8, 197–236.

van Wolkenten, M., Brosnan, S.F., and de Waal, F.B.M. (2007). Inequity responses in monkeys are modified by effort. *Proc. Natl Acad. Sci.* 104, 18854–18859.

Vonk, J., Brosnan, S.F., Silk, J.B. *et al*. (2008). Chimpanzees do not take advantage of very low cost opportunities to deliver food to unrelated group members. *Animal Behav.*.

Wallace, B., Cesarini, D., Lichtenstein, P., and Johannesson, M. (2007). Heritability of ultimatum game responder behavior. *Proc. Natl Acad. Sci.* 104, 15631–15634.

Walster, H.E., Walster, G.W., and Berscheid, E. (1978). *Equity: Theory and Research*. Boston, MA: Allyn and Bacon.

Warneken, F. and Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science* 311, 1301–1303.

Warneken, F., Hare, B., Melis, A.P. *et al*. (2007). Spontaneous altruism by chimpanzees and young children. *PLoS Biol.* 5, e184.

Whiten, A. (1998). Imitation of the sequential structure of actions by chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* 112, 270–281.

Wiesenfeld, B.M., Swann, W.B., Jr, Brockner, J., and Bartel, C.A. (2008). Is more fairness always preferred? Self-esteem moderates reactions to procedural justice. *Acad. Management J.* 50, 1235–1253.

Wynne, C.D.L. (2004). Fair refusal by capuchin monkeys. *Nature* 428, 140.

Xiao, E. and Houser, D. (2005). Emotional expression in human punishment behavior. *Proc. Natl Acad. Sci.* 102, 7398–7401.

Yamagishi, T. (2007). Reciprocity, strong reciprocity, and fairness. Paper presented at the Society for Experimental Social Psychology, Chicago, IL.

Zizzo, D.J. and Oswald, A. (2001). Are people willing to pay to reduce other's incomes? *Annales d'Economie et de Statistique*, 63–64, 39–62.

CHAPTER

20

Neuroeconomics of Charitable Giving and Philanthropy

Ulrich Mayr, William T. Harbaugh, and Dharol Tankersley

OUTLINE

INTRODUCTION

This chapter examines what neuroscience can tell us about people's motivations to pay for the provision of public goods and services that benefit society as a whole. The focus is on voluntary decisions about charitable giving. However, since most funding for public goods comes from taxation, we also consider the motives that lead people to vote for tax systems that compel themselves and others to pay jointly for public goods. We concentrate on research that adopts the usual economic assumptions of stable preferences and rational choices, but also consider the consequences of alternative, weaker assumptions.

The chapter starts with standard neoclassical utility theory and the implications of the assumption of selfish preferences, or what is rather confusingly called Pure Altruism. We show that the implications of this model are at odds with many real-world and laboratory behaviors involving voluntary contributions to public goods. Most glaringly, while this model predicts free-riding – benefiting from public goods without contributing to them – by all but a few donors, in reality charitable giving is common, if not quite universal. This suggests that the Pure Altruism model does not provide a complete account of behavior in public good situations – it does not accurately reflect the constraints that people face, or it does not accurately incorporate people's fundamental preferences, or people are making systematically irrational decisions.

Given this conflict, we then explore some alternatives to the Pure Altruism model, starting with the possibility that repeated interactions or attempts to signal one's wealth or character might explain contributions, before considering "warm-glow" models of altruism, where people derive a benefit from the act of giving. We show how these changes in constraints and preferences imply behaviors that are substantially different from those in the Pure Altruism model, and argue that these alternative models are useful for explaining a substantial amount of observed giving behavior.

The second part of the chapter looks at current neural evidence that addresses the reasons why people sometimes show altruistic behavior. After showing that the few studies that have directly dealt with motives indicate evidence of both Pure Altruism and warm-glow motives, we point to a number of open issues that deserve further study.

One question worth discussing at the outset is why neural evidence should be useful for testing models of public good decisions. One simple reason is that neural data provide a way of testing economic models that is very different from the empirical and experimental data that have typically been used. This presents the possibility of independent evidence in favor of, or against, various models of behavior. Another reason is that the neural evidence builds on a rich base of knowledge from animal physiology and from human brain-imaging work about the functional relevance of specific neuroanatomical areas. This means that there is the potential for generating interesting new hypotheses about motivations from observed patterns of neural activations. Moreover, results of neuroeconomics studies can feed back into the neuroscience literature, and further increase our general understanding of how the brain makes decisions and experiences their consequences.

Of course, these benefits are not specific to the domain of public good decisions, but hold for any economic question. More specific to public good decisions is the fact that neural data can provide information that it is very hard to obtain with other methods. Most existing economic studies on the reasons for giving are concerned with behavior, which provides only an indirect and one-dimensional window onto the mechanisms supporting altruistic motives. More direct ways of assessing motives, such as questionnaires, are often even less satisfying. People's self-reports on why they give – or don't – are suspect, since responses may reflect social demand effects, or the desire to report what they believe is socially acceptable or will please the experimenter, rather than their actual motives. In contrast, it should be quite hard for a person to "simulate" specific patterns of brain activity. Thus, neural evidence holds the promise of providing direct and physical evidence of the "hidden motives" behind altruistic behavior.

SOME ECONOMICS OF PUBLIC GOODS AND CHARITABLE GIVING

Introduction

We start by defining public goods and explaining why, given standard economic assumptions about people's preferences, private markets may not provide these goods adequately. After discussing how charitable giving and taxation can be used to provide these goods, we turn to the issue of voluntary contributions in more detail, focusing on the question of what motivates the contributions. This leads to discussion of how neuroscience can address those motivations.

Public Goods Defined

Most charity involves what economists call *public goods*. The definition of a public good begins with a classification system that distinguishes goods or services on two dimensions: whether or not one person's use of the good diminishes its value for others (rival or non-rival), and whether or not a person who doesn't help to pay for the good can be prevented from using it (excludable or non-excludable). Goods that are rival and excludable, such as food, housing, and medical care, are called *private goods*; goods that are both non-rival and non-excludable are called *public goods*. Classic examples of the latter are national defense, environmental quality, and knowledge. The fact that one person benefits from these things does not diminish their value to other people, and people can't be prevented from enjoying the benefits these goods produce.

Not all goods fit nicely into the private/pure public categories. Some goods pass one test but not the other. *Congestion goods* are rival, but non-excludable – for example, the use of a public road during rush hour. Since technological improvements tend to reduce the cost of exclusion, congestible goods can, with time, become private. Barbed wire is a classic example of a technology that made it feasible to create private property rights to grazing land. GPS toll-collecting technology may do the same with roads. *Club goods* are goods that are non-rival but excludable. Examples include concerts and classroom education – at least until there is crowding, one more user doesn't diminish the benefits others receive. However, people who don't pay can easily be kept from consuming these goods. Population increases tend to make non-rival goods rival – for example, as roads become crowded.

Aid to the poor, say for food and shelter, is another sort of good that doesn't fit perfectly into the classification scheme. Money spent to help a poor person creates a private good to the recipient; however, as Kolm (1969) points out, it also provides a non-rival and non-excludable benefit to all who want to live in a world where poor people are well fed and housed. Everyone who has this preference benefits from greater welfare expenditures, even if they haven't helped to pay for them. In this sense, the welfare of the poor is a non-rival and non-excludable public good.

Some voluntary giving is directed toward non-public goods. For example, monetary transfers between family members are common and often large. These gifts don't provide non-rival benefits (outside of other family members, at least), and are probably better studied using models that incorporate voluntary exchange and reciprocity. Similarly, gifts to education, the arts, and churches are often used to subsidize excludable goods. Gifts to provide congestible goods (such as scholarships to college) are also common. These are often thought of as charitable gifts as well, and to some extent most of the discussion in this chapter will apply.

Pure Altruism and the Optimal Level of Public Goods

Neoclassical economic models begin with the assumption that people choose from the available options so as to maximize their utility, or satisfaction. The simplest version of this that works for our purposes comes from Samuelson (1954). We start with the assumption that each member of society cares about his or her consumption of a private good x (usually thought of as total dollars spent on private consumption) and a public good G. We say that people maximize a utility function of the form $u_i = u_i(x_i, G)$ where i indexes each individual. Note that this is a model of Pure Altruism: by assumption, people care about how much G there is, but they do not care at all about how that level is achieved, except insofar as it affects their consumption of x – for example, when they make a contribution or pay a tax that reduces their x_i, but increases G. In other words, the standard model assumes that people care about the level of the public good only insofar as it contributes to their own consumption.

The fact that a public good is non-rival means that, in determining the socially optimal or efficient quantity of the good, we must consider the benefits that *everyone* will get from a given unit of the good. This is why G appears without a subscript and is in contrast to private goods, where only one person can consume the good and so only one person's benefit counts.

To keep things simple, we will assume two potential donors, person 1 and person 2, and measure utility in dollars. We will use the welfare of the poor as the public good; this benefits all people who care about the poor, so the benefit to "society" from an additional increment must include the benefits to both 1 and 2. (For simplicity, the private benefit to the poor is ignored here.) As a numerical example, assume that each person has an endowment m of \$5, and each unit of welfare for the poor of G costs \$4 to provide. Each unit of G benefits person 1 by \$3 and person 2 by \$6, so we are assuming that person 2 derives the larger benefit from seeing the poor taken care of. Each person can either keep all their \$5 or give \$4 to buy one unit of welfare. The total level of G is the sum of the amounts provided by persons 1 and 2, or $G = g_1 + g_2$. We could write these utility functions as $u_1 = x_1 + 3G$ and $u_2 = x_2 + 6G$, and the constraints as $x_i + g_i <= m_i$. Thus, person 1's utility is his starting \$5 minus \$4 if he donates, plus \$3 times the number of units of welfare G. Person 2's utility is his starting \$5 minus \$4 if he donates, plus \$6 times G. Each unit of the public good benefits both persons 1 and 2, so the combined utility for the two is given by $u_1 + u_2 = 5 + 5 - 4G + (3 + 6)G$. Obviously this is increasing G, since the cost of a unit of G is 4, while the combined benefit to persons 1 and 2 is $3 + 6 = 9$, so it's socially optimal to provide the maximum possible 2 units of G. This principle, that in finding the optimal level of G we need to weigh the costs against the benefit to *both* persons 1 and 2, follows directly from the non-rival nature of such goods, or from the fact that each unit brings benefits to more than one person. This optimality rule is known as the *Samuelson Condition*.

		Person 2	
		Keep	Give
Person 1	Keep	5, 5	8, 7
	Give	4, 11	7, 13

FIGURE 20.1 Modeling charitable giving as a strategic game. Each person starts with $5 and can buy one unit of the public good *G* for $4. Person 1 gets $3 in benefit from *G*, person 2 gets $6. Each cell shows the net payoffs to person 1, person 2 from that outcome. The total payoff is highest when both give, but person 1 does better for himself by keeping his money.

The Pure Altruism Model Implies Insufficient Levels of Public Goods

Since public goods are non-excludable, it is clear that a society that relies on voluntary action is going to have a problem in achieving the optimal level of *G* (Becker, 1974; Bergstrom *et al.*, 1986). A simple version of the argument using the example from above is shown in Figure 20.1. Each person can either keep all his $5, or give $4 to buy a unit of the public good. There are two choices for each of the two people, so four possible outcomes. The numbers in each cell show the payoff to person 1, then to person 2.

This situation is a modified version of the prisoners' dilemma from strategic game theory. Person 1 gets a higher payoff from keeping than giving, regardless of what 2 does, since $5 > 4$, and $8 > 7$. The twist is that person 2, because of the high benefit he gets from the public good, would prefer to give regardless of what person 1 does, since $7 > 5$ and $13 > 11$. In total, the two are better off if they both give, but only person 2 will do so voluntarily, thus we will end up with one unit of *G*, rather than two. Economists say that person 1 will *free-ride* on the donation of person 2, meaning that he will consume the *G* which 2 paid for, without making a contribution himself.

Mandatory Provision and Taxation

The usual solution to this free-riding and the inefficiently low level of *G* is some sort of political mechanism to determine how much *G* society wants, coupled with mandatory taxes. Funding public goods with taxes raises some issues, however. First, the process that determines the level of the public good is necessarily imperfect. People will generally have incentives to misrepresent their most desired level of the good if others will be helping to pay for it. Second, taxation generally means that even people who don't want the good have to pay for it, leading to arguments about fairness and coercion. Third, raising revenue by taxation is generally inefficient. Income tax, for example, distorts people's work effort away from the most efficient level. Last, governments and bureaucracies rarely have an incentive to provide public goods in the cheapest possible manner, since they don't get to keep the surplus that would result if they did.

Thus, the funding of public goods is a balancing act – both voluntary and involuntary mechanisms present serious problems. In general, modern societies rely much more on taxation than on voluntary giving to provide public goods. However, for specific goods (such as the arts, some kinds of medical research, some social welfare programs, and higher education) voluntary giving can be quite important. In the US, about 3% of GDP enters the "voluntary sector" through charitable giving. Interestingly, most other modern economies have much lower rates of giving.

Other Arguments Against the Pure Altruism Model

The free-riding result from the above model of Pure Altruism, where people care about the level of *G* and not how it is achieved, is straightforward and, perhaps unfortunately, very intuitive even to people without formal training in game theory. However, clearly not everyone always free-rides when given the opportunity, even when the benefits from another unit of *G* are less than the cost. In fact, by some measures, the amount of free-riding is shockingly low. Data from the Independent Sector estimate that, in 2003, 89% of US households gave something, with the average amount among those giving being $1620. Results in economic experiments designed to simulate the main aspects of contributions towards a public good also find large numbers of people giving substantial amounts of money (Isaac and Walker, 1988; Andreoni, 1995). In addition to free-riding, the Pure Altruism model makes other strong predictions about the voluntary provision of public goods. We start by examining those predictions, then discuss the evidence concerning them, before concluding this section with a discussion of alternative models of voluntary giving that are more consistent with the evidence.

Evidence About Who Gives and How People Respond to Changes in Levels of Public Goods

One interesting prediction of the Pure Altruism model is that government expenditures on a public good will "crowd out" or reduce private contributions towards that good – potentially, dollar for dollar (Bergstrom *et al.*, 1986). The logic for this is straightforward. Donors give until they just balance the extra utility they get from consuming the private good with the extra utility they get from consuming the public good. If the government takes a dollar in taxes from them and uses it to provide more of the public good, donors can get back to their original, optimal situation by reducing their giving by the dollar amount of the tax and spending the money on the private good. So, in theory, the only way government provision can increase the level of the public good is if the taxes are collected from non-donors. In practice, this is not what happens. Government expenditures do crowd out some private giving, but it's usually substantially less than dollar for dollar. Kingma and McClelland (1995) re-analyzed data from an earlier study involving donations to National Public Radio, and found almost no crowd-out effect.

Another significant challenge to this model comes from Andreoni (1989). He shows that, under very general assumptions (compared with the restrictive ones used in Figure 20.1), the Pure Altruism model predicts complete free-riding by all but a few people with the highest benefits from the public good, even in groups as small as 20. As the number of people benefiting from the good increases, the predicted number of donors shrinks to one, and the amount of the good provided stops growing. Significantly, this theoretical result holds even when people have very strong preferences in favor of the public good. A simple version of this can be seen in Figure 20.1. Person 1 cares quite a bit about the public good – but person 2 cares still more, and person 1 does not care enough to "top off" person 2's contribution. This shows the difficulty in using behavioral information to make inferences about preferences, when it comes to public goods.

Again, this predicted free-riding is not what actually happens. Participation in charitable giving is remarkably widespread, and even low-income people give, despite the fact that they would seem to benefit more from spending money on their own consumption than from the small increases in the level of public goods that their gifts can provide. In the US, for example, data from the independent sector for 1995 show that nearly half of households with incomes below $10,000 gave, with the average donation from those that gave exceeding $300 (Andreoni, 2006).

So, the Pure Altruism model fails three straightforward empirical tests: there is considerably less free-riding than the model predicts; people don't respond to government-financed changes in the amount of the public good by reducing their contributions dollar for dollar; and people who might not be expected to give, such as those with low incomes, actually do give. This does not mean that purely altruistic motives do not exist, but it does mean that we need something else to explain why people give their money and resources to charitable causes. Here, three broad classes of explanations are considered. First, we examine explanations that don't require altruistic preferences of any sort. Behavior might look altruistic because we don't understand the true constraints and structure of the game – what seems on the surface to be altruism might actually just be giving to signal wealth, or reciprocation in a repeated game. Next, we turn to the possibility that the Pure Altruism model of preferences is wrong and, in particular, consider models where the amount of a person's gift – as distinct from the level of the public good – is in the utility function. Both these explanations assume that people are rational. Finally, we consider the possibility that people are not rational, and turn to non-economic explanations for altruism.

Alternative Models of Altruism

The Pure Altruism model described above is sparse – maybe too sparse. One possibility is that the constraints or the rules of the game are unrealistic. For example, suppose that a reputation as a giver endears you to people and makes them willing to pay more for the products you sell. In this case, even a selfish person might give. The argument might be that in these circumstances it doesn't cost you a dollar to give a dollar to charity; instead, it actually pays off in more than a dollar in profits. This behavior is not really motivated by altruistic concerns, and could be modeled with a straightforward change in the budget constraints faced by the giver. Another possibility is that our model of preferences is wrong. Perhaps people not only care about the social problem, but also feel particularly good when they personally take voluntary action to help with the solution. As we will show, this can be modeled with a change in the utility function, as is done in "warm-glow" models of altruistic behavior.

In practice, economists sometimes use a "reduced-form" model to deal with the first possibility above – that our model of constraints is too simple. Reduced-form models subsume elements of the actual model structure into a simplified overall representation. This can be done by writing out a utility function that incorporates the additional constraints implicitly, using the argument that the behavioral predictions of the model are the same as would be obtained with simpler preferences and explicit constraints

(Harbaugh, 1998a, 1998b). But one of the major advantages of neuroeconomics is the ability to obtain direct readings of the neural activity associated with preferences, and so we argue that it is essential to be explicit about modeling changes in the conditions of the situation as changes in constraints, and likewise with changes in preferences.

Constraints

Charitable decisions typically involve repeated interactions between people who know each other. A basic result from game theory is that optimal behavior in a situation where people interact repeatedly can be very different from that in a one-shot setting. We can explore this by modeling charitable giving towards a public good as a simple game with n players. Assume that each player has m units of money, which they can keep or contribute to the public good. The sum of the individual contributions determines the level of G, and an individual's payoff is $m - g + \alpha G$. Note that this is a simple form of the Pure Altruism model $u = u(x, G)$ where $u = x + \alpha G$. By design, $\alpha < 1$ so it's optimal for each individual to give nothing, regardless of what the other players do, while $\alpha n > 1$, so the group does better if everyone gives all of their endowment. If $n = 2$ and α is set to, say, 2/3, this game is the classic prisoners' dilemma. (This set-up differs from the game in Figure 20.1 in that α is the same for everyone and is always less than 1.) In a one-shot game of this sort, *not* contributing is the optimal strategy for each player.

However, many giving decisions involve repeated decisions by people who can identify each other and who have some memory of what happened last time. Repeated interactions (with an indeterminate last round) change the game and the optimal strategy. Computer simulations with repetition show that Rapaport's tit-for-tat strategy (Axelrod, 1984), in which players start by contributing and then reciprocate their partner's previous choices, does substantially better at maximizing individual payouts than does the don't-contribute strategy.

Experiments with people, in either two-person or multi-person set-ups, produce very similar results. People frequently play tit-for-tat or variations on it, and the end result is considerably more voluntary giving – and higher individual payoffs – than the one-shot model predicts (see Andreoni *et al.*, 2008, for a review). We argue that giving in these settings should not be called altruistic, and we think that studies which incorporate neural evidence are well suited for further investigation of this issue, because they offer the potential for revealing distinct physiological mechanisms supporting giving behavior in situations where they stand to benefit from contraints of reciprocity versus in situations where there is no personal material gain to be derived from giving.

Repetition of the same game is not the only interesting situation involving complicated constraints. Another sort of model involves the possibility of incomplete information. For example, people might not know each other's ability, wealth, or preferences. You might only want to do business with people who are by nature somewhat altruistic, because they are less likely to cheat you, thereby creating a need to both acquire information about others, and signal information about yourself to potential partners. Talk is cheap, but charitable giving is a way to credibly signal your altruistic nature. Giving is cheaper the more you care about others, and this makes it hard for selfish people to mimic altruistic ones (Ariely *et al.*, 2007). A similar argument can be made about giving as a signal of wealth (Glazer and Konrad, 1996) or (for the arts) taste and sophistication.

Warm-glow and Impure Models of Altruism

The Pure Altruism model used so far assumes $u = u_i(x_i, G)$. While people care about the level of the public good, G, they don't care how that level is achieved. But people often say that "they want to make a difference," or that they get a good feeling from making a sacrifice that benefits others.

One way to incorporate this preference for a personal contribution to G is with the warm-glow model of giving (Andreoni, 1989, 1990). The most general form assumes preferences are given by $u = u(x_i, g_i, G)$. People still care about the overall level of the public good, G, but now they also care about their own contribution to the good, g_i. This is known as imPure Altruism, since it contains both the pure and the warm-glow motives. (Note that while Pure Altruismis the benefit a person gets from the level of the public good, warm-glow altruism is, in a certain sense elaborated on below, more selfish. If the public good is welfare for the poor, then warm-glow altruists give because of the good feeling *they* get from giving, while pure altruists give to make the poor better off.)

This apparently minor change in the assumptions about preferences has significant effects on giving behavior. In particular, none of the three empirical facts noted above – incomplete free-riding, frequent contributions from low-income/low-demand donors, and incomplete crowding out of private donations by government provision of a public good – are implied by the impure model of altruism. This follows from the assumption that charitable contributions provide a direct private benefit to the giver: the only way to get this benefit is to give. Tests using data on charitable giving (Ribar and Wilhelm, 2002) and from economic experiments (Palfrey and Prisbrey, 1997; Goeree *et al.*, 2002) provide support for this impure model, but with differing degrees of weight on the pure and the warm-glow elements. Later in the chapter, we show how neural data can also be used to test this model.

It seems strange to argue that people get a private benefit from giving up money, so the warm-glow model immediately brings to mind the question of where the preference for g_i originates. As mentioned above, one view of warm-glow giving is that it is essentially a simplified or "reduced-form" way of modeling the effect of repeated play or incomplete information. The argument is that people are primarily selfish, but they interact in an environment where reciprocation or signaling is rewarded. Rather than modeling the whole game explicitly, we just add g_i to the utility function. This is the argument Harbaugh (1998a) makes for including the "prestige motive" for giving, by which a person's well-being increases with the amount that other people believe he has given to charity, in utility, rather than modeling it as a way of increasing income. Other explanations for the warm-glow motive are discussed in the neuroscience section below.

In which model are donations less selfish? It's not obvious. Consider as extremes person 1 with Pure Altruism, $u_1(x_1, G)$, and person 2 with warm-glow altruism, $u_2(x_2, g_2)$. Both donate only because doing so increases their utility, and in that sense both types are selfish. Now consider a given donation – say, \$100 to fund a bridge. As Andreoni (1989) has shown, in a reasonably large economy a purely altruistic person would only give if he had a very strong preference for the good. So, this donation seems selfish in the ordinary sense of the word. Now consider person 2, the warm-glow giver. Again, he is only giving because he likes the feeling of giving. It seems difficult to establish that there is any difference in the degree of altruism the two people exhibit on the basis of this \$100 gift.

But now suppose that the public good G is the well-being of the poor. A \$100 donation from the Pure Altruism donor 1 will only occur if he has a very strong preference for the welfare of the poor – strong enough to override the tendency to free-ride. Such a donation certainly should be called selfless. And what about the warm-glow donor 2, who gives the same \$100 to aid the poor? This donation is motivated not by the donor's concern for the poor, but by his concern that he himself gives (or is seen to be giving) to improve the welfare of the poor. It seems entirely reasonable to argue that a purely altruistic donor who gives is "more altruistic" than a warm-glow giver who gives the same amount. ImPure Altruism would be somewhere in between. In short, while it seems difficult to label a person as more or less selfless on the basis of whether his preferences are defined by the Pure Altruism or the warm-glow model, it does seem reasonable to say that a given gift is more altruistic when it comes from a donor who is motivated by Pure Altruism. Ribar and Wilhelm (2002) address a related issue in the context of the Impure Altruism model and show that, even if people have a strong pure altruism motive, as societies get larger the donations we do observe will almost always be attributable to the warm-glow motive.

While at first glance it might seem unlikely that models based on utility maximization can be consistent with giving money away, we have shown that actually such a result is quite straightforward. All the above models make the standard economic assumption that people have well-defined preferences, and that they choose purposively to maximize their well-being given those preferences. We have shown that these assumptions lead to testable predictions, and that the results of empirical tests have generally supported the imPure Altruism model. However, all these tests leave substantial amounts of unexplained variance. In the next section, we will consider the counterargument to rational choice. Perhaps choices about when to give, and how much, are dominated by transient emotions and easily manipulable frames, or are so ingrained by morals or religion that changes in economic constraints have little effect.

NEURAL-LEVEL EVIDENCE

Neuroeconomic studies of altruism and charitable giving are just beginning. Here, we review the small set of works that deal directly with altruistic giving, and discuss a broader range of studies that have at least an indirect bearing on how experiments that produce neural data can help to answer the question, why do people give? Rather than reviewing the literature study by study, this section is organized in terms of what is known, and what neuroeconomics may be able to discover, regarding key questions. The focus is on the broad topics of whether or not giving decisions are rational, and what motivates giving.

Are Giving Decisions Really Rational Choices?

A basic assumption of all economic models of altruistic behavior is that giving decisions are rational responses to people's preferences and constraints. An alternative view is that while economic models may readily account for self-interested behavior, they cannot adequately explain behavior that is constrained by moral codes or that involves strong emotional responses (Greene *et al.*, 2001). However, we believe that the rational-choice model has held up well in the neuroeconomic tests that have been conducted so far.

First, we note that there is substantial evidence, collected from animals over decades, that specific brain areas – the ventral striatum and the nucleus accumbens, and (with some qualifications) the insula and the orbitofrontal cortex – respond to the fulfillment of basic needs such as food, shelter, and social contact. Functional imaging work has demonstrated that, in humans, these areas also respond to abstract rewards such as money (O'Doherty, 2004; Knutson and Cooper, 2005). New work is now beginning to extend these results to rewards in settings involving social exchange. King-Casas and colleagues (2005) studied the development of trust during reciprocal economic exchanges. They found that activity in the head of the caudate increased in response to benevolent transactions received from a partner. Because of animal and human work implicating the caudate in reward-based learning, the authors interpreted this caudate response as a neural reward signal. Later in the sequence of exchanges, enhanced caudate activity began to appear prior to transactions with that same partner. The authors interpreted this as an indication of an intention to commit a benevolent transaction. In other words, the caudate reward signal indicated a decision to give away money when there was an expectation of reward by a trusted opponent player. While this situation does involve gifts of money, it involves reciprocity, where the returns are private and do not have any public good component. As we argued earlier, optimal behavior in this situation could easily involve "giving" even by an entirely selfish person. Nevertheless, this study was among the first to show that neural activity in reward-processing mid-brain areas correlates with economic exchanges involving social behavior.

More direct evidence regarding the argument that decisions in charitable giving situations can be modeled as rational behavior comes from a study by Moll *et al*. (2006). These authors used fMRI to look at brain activity during decisions to give money to a range of different charities. They found increased activity in the ventral striatum both when subjects received money for themselves and when they chose to donate money, suggesting that overlapping areas are involved in computing the utility of receiving money and the utility involved with giving. Further, activity during costly donations in the ventral striatum and a nearby medial frontal area correlated with how much subjects actually gave. This result suggests that the same areas involved in computing the utility associated with receiving or losing money also register the utility resulting from donating money to help provide a public good. While this study does not provide evidence to distinguish between the warm-glow and Pure Altruism models of giving behavior, it does provide compelling evidence that giving decisions use some of the same basic preference computation mechanisms that are also used when no regard for others is implied.

A recent study by Harbaugh *et al*. (2007) looks further into these issues. These authors used a protocol that included a mandatory, tax-like condition in which subjects simply observed more or less costly transactions from their own account to the charity. They also included trials in which the charity received money at no cost to the subject, and trials in which the subject received money at no cost to the charity. They found increased activity in the head of the caudate and the nucleus accumbens when subjects received money for themselves, replicating the results discussed above. A novel finding, however, was that there was also increased activity in these regions when the subjects simply observed the charity receiving money. As further evidence of the importance of these regions for charitable giving decisions, Harbaugh and colleagues showed that the degree of activation in the mandatory condition can be used to predict voluntary giving. The logic for this follows directly from the Pure Altruism model of giving, where $u = u(x, G)$, and is illustrated in Figure 20.2. Here, x denotes money going to the individual subject, and G denotes money going to the public good. The lines show representative *indifference curves*, or combinations of x and G that yield the same utility. The slope of the curves is a measure of a person's willingness to give up one good for the other, which in turn is determined by the relative *marginal utility*, or relative contributions to utility from increased amounts of the goods. A selfish person, with a low (but still positive) marginal utility of G, would only give up a small amount of x in return for a given increase in G, while a more altruistic person would give up a larger amount of x for the same increase in G, giving him steeper indifference curves. Movement towards the upper right – increasing consumption

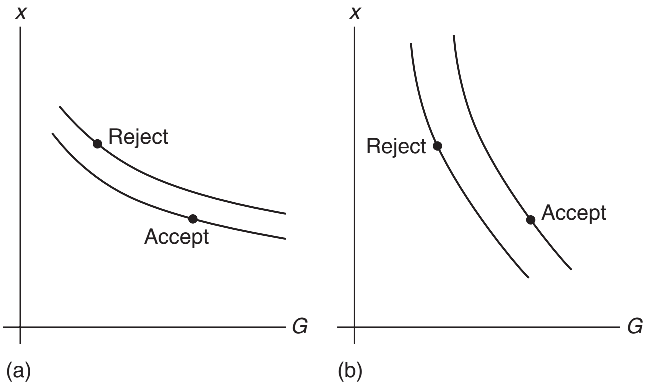


FIGURE 20.2 Altruistic preferences and decisions to accept or reject a charitable transfer: (a) egoistic; (b) altruistic.

of both goods – would put people on an indifference curve representing a higher level of utility. As Figure 20.2 shows, egoists get higher utility by rejecting transfers that reduce their *x* and increase *G* (i.e., they have lower marginal utility for *G*), while altruists maximize their utility by accepting them (i.e., they have higher marginal utility for *G*).

It emerges that the predictions of this model are supported by the combination of neural and behavioral data from the Harbaugh *et al*. experiment. The authors used increases in neural activity in the ventral striatum and insula during pure mandatory payments to the subject as an indicator of the subjective marginal utility of money (the private good *x*), and activity in pure mandatory payments to the charity as an indicator of the marginal utility associated with increases in *G*. In other words, the magnitude of activation in the ventral striatum in response to receiving money for oneself was taken as the measure of the person's marginal utility for the private good, and the magnitude of activation in response to seeing increases in the money available to the charity was taken as the measure of their marginal utility for *G*. The Pure Altruism model holds that the difference between these measures predicts people's willingness to give. As Figure 20.3 shows, people with large neural responses to *G* relative to their responses to *x* were particularly willing to give away money to the charity in the voluntary parts of the experiment.

Both the Moll *et al*. and the Harbaugh *et al*. studies support the argument that the neural basis of charitable giving decisions is consistent with a rational choice model where people make their giving decisions by comparing the utility they get from spending money on themselves with the utility they get from seeing the charity have more resources to devote to the public good. An open question, however, is whether Pure Altruism or warm-glow provides a better model of the utility of charitable giving. We turn to this question about the motives of giving in the following section.

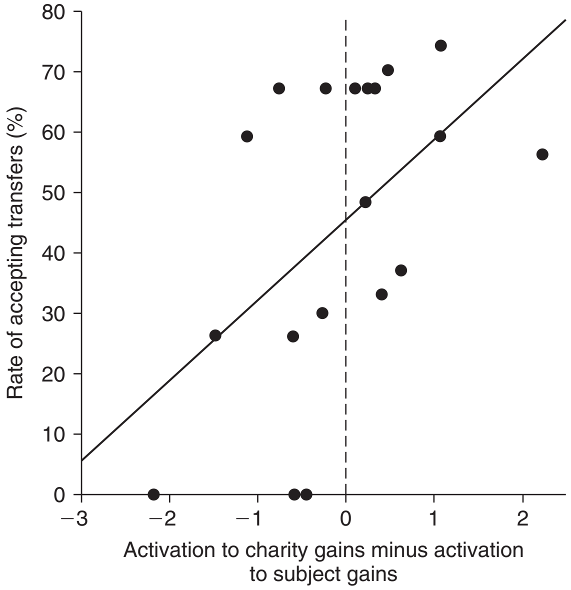


FIGURE 20.3 Neural activity in the mandatory condition predicts voluntary giving. Each point represents one subject. Negative amounts on the horizontal axis indicate that activation in the ventral striatum and insula is larger in response to monetary gains for the subject (*x* in Figure 20.2) than to gains for the charity (*G* in Figure 20.2). People on the right or positive side of the dotted line are more "altruistic," and have steeper indifference curves. The vertical axis shows the proportion of costly transfers that the subject accepted. The neural activation data support the Pure Altruism model illustrated in Figure 20.2.

However, it is also important to note that the existing evidence in favor of utility computations during giving decisions tells us little about the potential limits of the rational choice model. For example, it is at least conceivable that, in a situation for which strong moral codes exist, utility computations can be replaced by a purely rule-based decision process (e.g., "Always give away 10% of your income"). Adherence to such a rule might lead to behavior that is inconsistent with the utility computations for a particular situation. Alternatively, such moral codes might simply feed into the utility computations – for example, by linking rule-incongruent options with a high negative utility. Neural data that trace activity in areas related to goal and rule representations should help to uncover the potential interplay between rule representations and utility computations. More generally, altruistic decisions may be subject to framing effects, or contextual influences on decisions. For example, Andreoni (1995) provides experimental evidence that people end up allocating more money to a public good when their decision is described as taking money from the group than when it is described as giving. Such differences could be interpreted as irrational, or just as evidence that people rationally use "rules of thumb" to reduce the cost of their choices.

Although neuroscience has not yet addressed the influence of framing effects on charitable giving, several studies have begun to delineate the brain systems responsible for framing effects in moral and social judgment. Greene *et al*. (2001) asked subjects to judge the appropriateness of an action in various moral dilemmas. For example, in one dilemma a trolley is running down a track on which it will kill five people, and subjects must decide whether or not to pull a switch which will re-route the trolley, saving those five people but killing a single person on the other track. Subjects generally respond that it is appropriate to push the button and save the five people at the cost of the one. However, in a slightly different dilemma, subjects must decide whether it is appropriate to push a large person onto the track in order to stop the trolley and save the five people. Despite killing and saving the same number of people in each of the two dilemmas, subjects are significantly less likely to say that pushing the person onto the track is appropriate behavior. Greene and colleagues proposed that the difference in the two scenarios is that the pushing scenario activates emotional systems that are not active in the switch scenario, and that these emotional differences alter people's judgments. The authors found neural evidence to support this theory, insofar as brain regions associated with emotional processing, including posterior cingulate cortex, medial frontal lobes, and posterior parietal lobes, were more active during "personal" moral dilemmas – dilemmas involving emotionally intense behaviors such as direct personal harm. Further evidence for the importance of emotions as determinants of the sorts of social decisions that are involved in charitable giving comes from Koenigs *et al*. (2007), who found that lesions to the orbitofrontal cortex – a region associated with processing of emotions – were associated with an increased tendency to use "rational" utilitarian judgments in decisions involving tradeoffs between one person's life and another's.

Another study, by Delgado *et al*. (2005), suggests a similar framing effect in social judgments. Subjects read biographies of "good," "bad," and "neutral" fictional characters, and then played a trust game with the three characters. Despite the fact that all three characters cooperated at the same level, subjects throughout the experiment continued to trust the "good" partner and distrust the "bad" partner. Further, brain regions involved in reward-learning did not respond differentially to cooperative and defective behaviors of these "good" and "bad" partners, suggesting that prior social information reduces neural responsiveness to good and bad outcomes. Because charitable giving is likely to involve many of the same computations and decision-making mechanisms that support social and moral judgments, charitable behavior is likely to be subject to similar types of emotional and learning biases. We note, however, that not all these biases are irrational – in fact, in the context of repeated interactions discussed in the economics section, many of these biases are likely to represent optimal responses.

Can Neural Evidence Reveal the Motives for Giving?

If we agree that decisions to make charitable donations are at least in part rational, then we can consider the correct way to model the behavioral and neural mechanisms influencing such choices. As explained earlier, economic models have distinguished between two broad classes of motives, one referred to as Pure Altruism and the other as warm-glow, with Impure Altruism models including both. In the context of contributions to charities that provide benefits to the poor, Pure Altruism implies that the donor's utility of giving is directly linked to the utility of the poor. In contrast, a cynical view would be that, under the warm-glow motive, the recipient's well-being is a means to benefit the donor, by signaling wealth, showing off his character as a good person to others, or reducing guilt and making it easier to think about himself as a moral person. A key distinction between these two groups of motives is that a person only gets warm-glow benefits if he makes an active decision to give, while a purely altruistic motive should be satisfied even by passively observing an increase in the public good which is paid for by someone else.

The fMRI study by Harbaugh *et al*. (2007), cited above, contrasted exactly these two situations. In the mandatory trials, the transfers from the subject to the charity were done in a passive, tax-like manner. In the voluntary trials, the subject could make deliberate giving decisions. The fact that reward-center activity occurred even when the charity received money in the form of a mandatory tax provides the first neural evidence showing that exogenous changes in the amount of money a charity receives have a neural effect. However, this still leaves the question: to what extent can this neural evidence be considered a motive that influences actual giving decisions? As discussed above, Harbaugh and colleagues showed that the degree to which reward-center activity

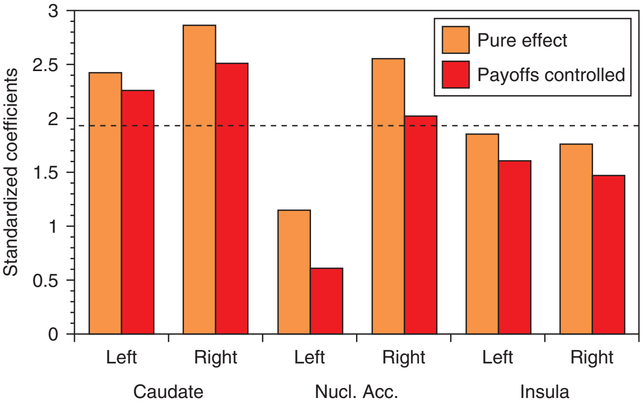


FIGURE 20.4 Higher activation when transfers are voluntary. The figure shows that activation in the indicated areas is generally higher in voluntary than in mandatory transfers.

responded to the pure, mandatory gains by the subject and the charity predicted the participant's "out of treatment" willingness to give in a voluntary manner. This suggests that a Pure Altruism motive exists and influences giving decisions (see also Batson, 1988).

Of course, this pattern of results does not rule out an additional warm-glow motive. As discussed in the economics section, the observed facts of charitable giving behavior show that there must be something more than Pure Altruism behind charity, such as a warm-glow motive. The design in the Harbaugh *et al.* (2007) study provides a way of distinguishing between the two motives, by comparing activity in areas that respond to reward – specifically, the nucleus accumbens, caudate, and insula – in mandatory and voluntary giving situations. As Figure 20.4 shows, activity in these areas was higher in the voluntary treatments. Importantly, the extra neural activation from voluntary giving was still present after controlling for the actual payoffs (i.e., the fact that in voluntary transfers individuals could keep more money for themselves than in mandatory transfers). The fact that the same monetary transfers are more rewarding when subjects can take responsibility for them shows that giving is not only motivated by concern for the recipient, as in Pure Altruism, but also by a person's desire to feel responsible for helping the recipient. This second motive is warm-glow altruism (or, as discussed in the economics section, a more complicated version of constraints, which can be reduced to warm-glow altruism). While the warm-glow model does not specify exactly how a person's voluntary giving enters utility, perhaps the most reasonable argument would be that utility increases with the amount of the gift. Harbaugh and colleagues (unpublished data) were only able to identify an increase in activation during voluntary giving. They did not show that activation increased with the amount given, and this remains a subject for future work.

While the Moll *et al.* and Harbaugh *et al.* studies produced evidence for the existence of Pure Altruism and warm-glow motives, it is clear that additional work regarding the neural basis of warm-glow giving may answer some intriguing questions. For example, one issue involves to what extent the higher activation during voluntary giving is tied to decisions about giving, or whether it is a specific example of the general "free to choose" principle, that more choices make people better off. Another issue concerns the strength of the evidence. While the Harbaugh *et al.* (2007) study found support for the warm-glow effect by demonstrating higher neural activity in the voluntary condition, a more rigorous test would be to demonstrate that activity increases with the payoff to the charity at a faster rate for voluntary than for mandatory transfers. Harbaugh and colleagues could not establish this pattern in a statistically reliable manner in their neural data, although this is perhaps not so surprising, given the small number of subjects. However, another possibility is that the contribution of warm-glow to utility is not always strictly increasing – perhaps people feel that a small contribution is enough for them to think that they have "done their bit."

It also might be the case that, in the Harbaugh *et al.* study, the warm-glow effect was reduced by the minimal nature of the manipulation. In particular, all transfers were anonymous. While even anonymous giving decisions provide some opportunity for the warm-glow feeling (i.e., one can feel good about oneself), outside the lab charitable giving and philanthropy are anything but anonymous – United Way contributions are recorded by employers, supporters of the Opera get their name or even their picture in the program, and people who make large enough contributions to a university can contract to have a building (or at least a brick) named after them.

This raises the point that neither Pure Altruism nor warm-glow are simple one-dimensional motives, and that much research can be done to unpack their constituent components. We believe that neuroeconomic experiments are particularly well suited to this, as is discussed below.

Neural Evidence Regarding the Role of Empathy and Emotion

Evidence from social psychology suggests that when people are in more empathetic, sympathetic, or positive states, they are more likely to give money or volunteer time to charities or needy people.

Batson and colleagues have shown in a number of experiments that inducing empathy in subjects, for example by instructing them to think about how another person feels, increases the rate at which subjects are willing to help another person (Batson *et al.*, 1988; Batson, 1991).

The role of empathetic states in giving behavior offers a potential explanation for the identifiable victim effect, in which altruistic behavior is strongly influenced by whether or not givers are exposed to an individual recipient in need (such as a single hungry child). Slovic (2007) shows that presenting multiple victims (starting with two!) usually reduces the altruistic response to a humanitarian disaster. One possible interpretation of these results is that the system that tracks others' bodily or mental states has a severe capacity limitation (Morewedge and Schooler, 2008). Researchers such as Slovic argue that stimuli with more tangible properties, such as photos of a suffering individual, are more capable of engaging emotional systems and thereby evoking the affective states that lead to altruistic behavior than are abstract stimuli such as death-rate statistics. The finding that the presentation of a single face significantly increases donations supports this line of reasoning, since faces are known to be highly arousing and emotional stimuli (Adolphs, 2002).

To the extent that altruistic behavior requires the ability to recognize the needs of other people, we should expect altruism to depend upon the functional integrity of brain systems required for such social perception capacities. Recent neural work by Singer *et al.* (2004; see also Chapter 17 of this volume for additional details) and others has mapped out the neural circuitry associated with empathic responses to observing physical harm done to another. They find that the anterior insula seems to function as a representational system that traces one's own and others' bodily states. Interestingly, neural activity in this region while observing somebody else receiving physical harm showed a considerable correlation with a self-report empathy scale. More cognitive forms of empathy in which subjects are instructed to think about the feelings of another person have implicated a prefrontal-parietal set of brain regions that is thought to be involved in understanding the intentions and predicting the behaviors of other agents (Ruby and Decety, 2004).

Despite this, so far there is no direct neural evidence linking empathy to altruistic behavior. In fact, Harbaugh *et al.* (2007) found no correlation between psychometric empathy scales and either behavioral or neural manifestations of altruistic tendencies (unpublished data), and Tankersley *et al.* (2007) found no correlation between empathy measures and the neural responses that predicted altruistic tendencies. However, aside from the fact that such null-findings are not very conclusive, it is also important to note that the giving situation implemented in these studies was relatively abstract, and did not involve identifiable recipients with specific characteristics or needs.

The Tankersley *et al.* study does suggest, however, that the neural systems that are involved in social perception might serve as precursor functions to empathy. In their experiments, Tankersley and colleagues measured brain activation while subjects either played a simple game or watched as a computer played the game. After scanning, subjects completed a self-report altruism survey. The authors found that the posterior cingulate cortex and bilateral posterior parietal lobes were more active when subjects watched the computer play than when subjects played the game themselves. Interestingly, activity in the right superior temporal cortex was correlated with altruism. More altruistic subjects (as measured by self-reports about behavior outside the lab) showed higher levels of activity in this region, in comparison with less altruistic subjects. Although they did not find significant correlations between measures of empathy and activity in the right superior temporal cortex, this region is thought to be involved in a variety of social perception tasks, including the detection of agency and goals. Indeed, Blair (2005) has proposed a model in which this region supports early analysis of actions, which are in turn transmitted to other systems that facilitate empathic responding.

The Tankersley *et al.* study shows the contribution that neuroscience can make towards understanding the components of altruistic motivations. In their paradigm, right superior temporal cortex activated in response to the flash of a target, indicating that the computer had responded for that trial. The fact that this region responded more to such a simple stimulus for altruists than others suggests that fairly low-level attentional mechanisms are influencing subjects' propensity to engage in helping behavior. Because neuroscientific techniques offer access to motivational mechanisms of which subjects are unaware, they are uniquely suited to reveal findings that individual differences in visual attention influence altruistic tendencies.

In addition to empathy and its perceptual precursors, social psychologists have also shown that simply inducing positive affect – for example, by arranging an experimental setting where the subject "finds" a dime in the phone booth (Levin and Isen, 1975) – increases helping behavior (mailing a stamped addressed letter "accidentally" left at the phone booth). The authors speculated that the increased willingness to help was an effort to maintain the good mood induced by finding the dime. Although this inference might seem far-fetched, a wealth of psychobiological data supports a connection between prosocial behavior and a variety of positive affective states or dispositions.

Depue and Morrone-Strupinsky (2005) use the term *affiliation* to describe an individual's motivation towards social stimuli, as well as their capacity to enjoy social stimuli. Affiliation is a complex trait related to other traits such as extraversion and sympathy, which are in turn predictive of prosocial behaviors, including altruistic behavior. Depue and Morrone-Strupinsky argue that the dopaminergic system, including the substantia nigra and its subcortical (basal ganglia) and cortical (medial prefrontal cortex) projections, influences an individual's sensitivity and motivation towards social rewards such as friendly vocalizations or facial expressions, as well as a wide range of positive affective traits, from desire to self-efficacy. They argue that, in this case, a second opioid-based system is critically involved in experiencing the pleasure of these social stimuli, and that the dopaminergic and opioidergic systems interact to produce learning and memory of social stimuli and their rewarding values. Further, research suggests that genetic differences significantly influence the functioning of the opioid and dopamine systems, and these genetic differences have been related to significant differences in motivation, including prosocial motivation, in animals (Depue and Collins, 1999).

Recent work on the effects of the neuropeptide oxytocin actually provides some indirect evidence that empathy leads to altruistic behavior (Zak *et al.*, 2007). Participants who received either oxytocin or a placebo played both a dictator game (i.e., where the player can split money with an anonymous recipient with no opportunity for reciprocal punishment) and an ultimatum game (i.e., where the recipient can punish the giver for unfair sharing). Subjects who had received oxytocin made substantially large offers, but only in the ultimatum game and not in the dictator game. The authors argue that only the ultimatum game requires emphatic appraisal of the others' mental states. They therefore conclude that oxytocin, which is known to promote bonding and affiliative tendencies, boosts altruistic behavior by increasing empathy, but not by affecting altruism *per se*. The somewhat puzzling aspect is that in this case greater altruism after receiving oxytocin is observed under conditions of strict reciprocity – not a situation that is typically regarded as a good test of Pure Altruism. Thus, it would be interesting to replicate these results with other ways to manipulate demands on empathy. In this context, another recent study (Knafo *et al.*, 2007) is of interest; this looked at people who differed in the promoter region of a gene that is involved in the receptor for vasopressin, which chemically is very similar to oxytocin (AVPR1a). In a one-shot dictator game, individuals with long versions of the relevant gene gave substantially more money to the anonymous other player than did those individuals with the short version. Thus there is converging evidence about the relationship between biological mechanisms of social bonding and altruistic behavior in economic games, including initial findings that point to important genetic factors. However, the exact pathway and conditions under which such effects become apparent needs further study.

Aside from the propensity for empathy, another trait that is known to increase altruistic behavior is the complex trait of "resilience," which involves (among other things) the ability to respond in an effective manner to stressful environments, and to adapt to new social situations (Charney, 2004). According to Charney, several of the brain's neurochemical systems, including the dopamine system, influence altruistic behavior by regulating an individual's sensitivity to positive and aversive stimuli in the environment, which in turn influence the individual's affective and motivational state. Thus, there seems to be a cyclical effect in which helping behavior induces positive affect and resilience, which in turn increases the propensity to help.

We know from economic data and experiments that there are large individual differences in charitable giving. Only part of the variation can be explained by differences in demographic variables such as gender, income, or education. Rather, it seems that a substantial part of the remaining variance can be explained by potentially stable individual differences in aspects such as the propensity to show empathy, other-regarding emotions, or hormonal factors. Already, there is evidence that differences between people in altruistic behavior go along with interesting neural-level differences. Future work of this kind should reveal the neural and psychological basis of the individual characteristics that are critical during public good decisions.

Fairness as a Motive for Giving

The Pure Altruism motive in particular might be expected to respond to people's appreciation of the needs of other specific individuals or groups. There is evidence from psychological research that many people entertain a "just world" belief (see, for example, Lerner, 1980; Hafer and Begue, 2005). The experience of innocent victims endangers this belief, and can lead to altruistic behavior in order to reconstitute the "just world." Economists have proposed so-called altruistic punishment as a way of discouraging defectors and free-riders in the context of cooperative public good situations. In fact, laboratory studies have shown that people are often willing to pay a cost in order to punish unfairness – and indeed to reward fairness (Andreoni *et al.*, 2003; see chapter by E. Fehr in this handbook). The fact that this occurs even in non-reciprocal, anonymous situations suggests that punishment is driven by the desire for fairness itself, not by the prospect of later positive consequences for the punisher. In fact, neural evidence suggests that altruistic punishment activates very similar reward activity to that found in actual giving situations, and the strength of this activation correlates with the actual amount invested in order to punish defectors (de Quervain *et al.*, 2004).

Empirically, it may be difficult to distinguish between Pure Altruism that is driven by the perception that somebody's needs are fulfilled rather than by the abstract goal of establishing fairness across individuals. However, one interesting question is, to what degree is the willingness to engage in altruistic punishment correlated with the neural Pure Altruism Indicator established by Harbaugh *et al.*? Do Pure Altruism and altruistic punishment go together, or are there alternative ways to promote the provision of public goods? Maybe some people deriving utility from acting altruistically while others derive from punishing those who don't?

Goals and Altruistic Decisions

Computations of utility related to self-interest and other-interest need to integrate information from a variety of sources. One possibly important source is an individual's abstract goals and values. There is indirect neural-level evidence that goal representations can play an important role during altruistic decisions. Moll and colleagues (2006) found that activity in a small, anterior medial frontal region, in situations that pitted self-interest against other-interest, was highly predictive of self-reported engagement in charitable giving. According to the authors' interpretation, this prefrontal area represents and maintains high-level goals. By this logic, only individuals who can establish firm goal representations can withstand the pull towards the self-serving option of keeping money for oneself (see also related work by Knoch *et al.*, 2006).

An interesting point raised by this result regards the extent to which goal representations in the prefrontal cortex actually influence decisions, by feeding information into the utility computation. In this case, we should see a tight linkage between prefrontal goal activity and Pure Altruism reward-center activity. The alternative possibility is that representations of rules, goals, or moral codes represented in the prefrontal areas allow for direct control of behavior, bypassing the utility-based decision-making system. From a neuroeconomic perspective, this brings us to the fundamental question regarding the potential limits of the utility-based rational-choice model. In principle, it should be possible to address this question by creating situations in which the relative contributions of prefrontal areas and reward-processing areas in the midbrain to actual altruistic decisions can be determined. If prefrontal goal representations influence behavior by bypassing utility computations, then we should see activity in midbrain areas involved in reward processing and prefrontal activity as independent predictors, both for decisions across individuals and for variability across decisions within individuals.

Warm-glow, Private Versus Public Giving, and the Effect of Religion

Recent economic work on voluntary giving decisions has focused on establishing facets of the warm-glow motive (see above). This work has demonstrated that the prospect of signaling "good character" through one's charitable contributions is an important motivator of seemingly altruistic behavior, as in Harbaugh (1998a, 1998b) and Ariely *et al.* (2007). Such motives should be particularly aroused when the charitable behavior is public. Interestingly, recent psychological work provides at least indirect evidence that it takes only very little to create a "public" situation. Bateson *et al.* (2006) showed that a poster depicting a pair of eyes led to an almost three-fold increase in prosocial behavior (payments into a shared coffee fund) compared with a control situation involving a picture of flowers. While the cause is rather trivial, this is a dramatic effect.

So far, there is no neural-level work that attempts to separate the different aspects of warm-glow. If we go by the working hypothesis that midbrain reward areas are generally involved in utility computations, we would not necessarily expect that warm-glow and Pure Altruism can be distinguished within these areas. After all, both motives can be captured within a utility-based rational-choice model. In fact, Harbaugh *et al.* (2007) found that mandatory transfers to the charity (i.e., Pure Altruism) produced midbrain reward-center activity, and that voluntary transfers (i.e., warm-glow giving) produced additional activity in the same areas.

However, warm-glow giving might very well express itself in a distinctive manner in other neural areas. Most, but not necessarily all, aspects of warm-glow giving involve the consideration of an audience and repeated interactions. Specifically, prestige and signaling effects require that the potential donors have a mental model of the observers that represents what these observers might classify as financially impressive or morally admirable actions. Therefore, one might expect that a signature of warm-glow will be activity in "social cognition areas" such as the anterior rostal, medial prefrontal cortex (including paracingulate gyrus), which has been associated with the consideration of others' mental states (see, for example, Amodio and Frith 2006). So it is interesting to consider to what degree private and public giving lead to differing activity in these prefrontal areas. Complicating matters empirically, similar areas have been implied in the representation of one's own value judgments (e.g., Zysset *et al*. 2002). As a theoretical complication, it is also worth considering the so-called "objective self-awareness" theory (Duval and Wicklund, 1972). Supported by a host of empirical evidence (Silvia and Duval, 2001), this theory states that cues that focus people's attention on themselves (e.g., mirrors, cameras, or onlookers) also bring their own standards into the foreground. From a theoretical perspective, it is a significant difference whether somebody shows an increase in prosocial behavior in order to impress others or because the presence of others brings one's own standards into the foreground. It is too early to tell whether neural evidence can clearly distinguish between the process of activating one's own goals or standards, and the process of thinking about somebody else's mental states. Clearly, such evidence could be very useful to pinpoint the various ways in which private and public altruistic decisions might differ.

One of the most interesting puzzles in charitable giving is the effect of religiosity. In the US, nearly 60% of all giving goes to religious organizations (data from the Independent Sector, analyzed in Andreoni, 2006). In a review of the empirical literature, Bekker and Wiepking (2007) report that "Positive relations of church membership and/or the frequency of church attendance with both secular and religious philanthropy appear in almost any article in which this relation was studied." But they then note that experimental studies, e.g., Eckel and Grossman (2004), do not support this effect for giving to secular charities. Evolutionary theorists have speculated that the promotion of cooperative, prosocial behavior may be the most critical advantage brought by the evolution of a mental capacity for religious thought (Sosis and Alcorta, 2003; Wilson, 2002). Maybe the most straightforward explanation of this phenomenon is that religions typically establish strong standards and goals for giving, as in calls for tithing (Lansdell, 1906). The result in Moll *et al*. (2006) regarding the correlation between goal-area activity and actual giving might be an interesting starting point for looking into this possibility. However, what is so far largely missing is experimental evidence that allows us to establish (1) that religious beliefs and practices actually change constraints and preferences, rather than people with certain preferences selecting into religion, and (2) how religion might change preferences. The few experimental studies that do exist suggest that the evocation of a higher power leads to more prosocial behavior (for example, the mention of the ghost of a dead graduate student haunting the lab increases prosocial behavior; Bering *et al*., 2005). One possible explanation of such results is that a higher power functions as an all-seeing observer. We have already seen that even subtle cues of being observed can have powerful effects on prosocial behavior (Bateson *et al*., 2006). In a certain sense, decisions by religious individuals are always "observed," and so charitable decisions might always involve a larger warm-glow benefit. This hypothesis would suggest that religious individuals, or individuals in which a "god concept" was induced experimentally (Norenzayan and Shariff 2007), might consider how their own decisions are perceived by the "higher power." This leads to the prediction that religious individuals in private-giving situations and non-religious individuals in public giving situations should exhibit similar activity in mentalizing areas (Amodio and Frith 2006).

The Neural Basis of Life-span Differences in Altruistic Behavior

Individuals differ in their altruistic tendencies and, as discussed in the previous sections, the few existing brain-imaging studies have produced interesting insights into the nature of these differences (Moll *et al*., 2006; Harbaugh *et al*., 2007; Tankersley *et al*., 2007). This raises the hope that neural evidence can also be used to explain one of the most powerful predictors of individual's altruistic tendencies – namely, that the percentage of income devoted to charitable giving increases relatively steadily across the life-span (Andreoni, 2006). Since older people are typically wealthier people, this means that the majority of the donations that charitable organizations receive come from older donors.

So far, there is no explanation for this intriguing life-span trend. In fact, based on what we know about

(1) the negative effects of aging on frontal cortex (see, for example, Raz, 2005) and (2) the involvement of prefrontal areas in promoting behavior that goes against immediate self-interest (e.g, Moll *et al.*, 2006), it is a puzzle why altruistic tendencies seem to increase rather than decrease. In this context, it would be very revealing to examine prefrontal contributions during altruistic decisions across the life-span. It is very possible that there is actually a weakening of prefrontal influences that would lead to a reduction of altruistic behavior if it were not for additional factors.

What might such additional factors be? Interestingly, there is evidence that, as people's time horizon shrinks, they experience a gradual shift in general motivational orientation, away from information-gathering and achievement and towards meaningful emotional experiences – in particular in the social domain (see, for example, Carstensen, 2006). One aspect of this shift seems to be less regard for potential losses, as indicated by a reduction in loss-related neural responses (Larkin *et al.*, 2007). While so far this general perspective has not been applied to public good decisions, there does seem to be a natural connection. In the context of public good decisions, a greater tendency to engage in "meaningful interactions with the world" and a reduced appreciation of losses to oneself might lead to a greater Pure Altruism Response in reward areas.

More likely than not, the story of life-span effects on altruistic giving will be a complicated one. It will probably involve both prefrontal and reward-area contributions that represent the negative cognitive and the more positive socio-emotional effects of aging. Neuroimaging evidence should be useful in disentangling these counteracting forces.

CONCLUSION

We started this chapter by explaining the primary economic models of charitable giving: Pure Altruism and warm-glow altruism. We showed that the bulk of the evidence from empirical and experimental studies supported the existence of a purely altruistic motive for charitable giving, but that in societies of more than 20 or so people this motive lost its force, and could not explain the widespread giving that we observe. We explained the empirical, experimental, and neuroscience studies showing that the combination of Pure Altruism and warm-glow known as Impure Altruism was an appealing alternative model, and also why alternative explanations based on repeated play, reciprocity, and signaling might too be consistent with the data.

The above summary suggests a consensus which might indicate that the marginal benefit from further investigations is low. This would be a mistake. There is a plethora of interesting and important questions, and neuroeconomic techniques are well suited to investigating many of them. In particular, in the sections above we have discussed how neuroeconomic methods can be used to address questions about the determinants and components of the Pure Altruism and warm-glow motives.

There are additional questions to consider. While the focus of this chapter has been on charitable giving, the truth is that public goods are mostly funded by taxation. Taxation raises a set of issues that has not yet been addressed by neuroeconomics. Tax compliance is one example. Most economic studies show that tax cheating is far lower than would be expected if people simply weighed the benefits against the small chances of getting caught and the low subsequent penalties. Thus it seems likely that people pay taxes at least partly out of a sense of obligation, if not exactly altruism. Are the motives involved in tax compliance similar to or different from those in charitable giving and philanthropy? The fact that Harbaugh and colleagues found activity in reward-related areas such as the nucleus accumbens during tax-like transfers suggests some degree of similarity between motives involved when paying taxes and during charitable decisions. However, to what degree this generalizes to more realistic situations is an open issue that can be addressed with analytic techniques similar to those used by Harbaugh and colleagues.

A second and related set of questions involves voting for taxes. Assuming vigorous enforcement, paying taxes is not exactly altruistic. But what about voting to increase taxes that fund public goods? Taxes impose costs on the voters, and in that sense voting for a tax might sometimes be considered altruistic. However, taxation also imposes costs on others. In fact, it is the costs imposed on other people that will fund most of any increase in the public good. How do voters weigh these factors? When preferences on a given public good vary widely, so that some of those paying for a public good will not benefit from it (such as in the case of reductions in timber-cutting to save an endangered species), is the altruistic benefit of voting diluted?

Another interesting question is whether warm-glow benefits should be included in calculations of social welfare. If they are, we might argue that societies should rely more on philanthropy, and less on taxation. The logic would be that more voluntary giving – perhaps promoted by increased tax deductions – would make people better off both through the Pure Altruism preference for more public goods, and through the warm-glow preference for active giving. This would seem supportable if warm-glow motives arise out of a basic preference for taking action to help provide public goods. It might also hold if the warm-glow motive originates with the need to signal one's trustworthy character to potential partners. But suppose that the warm-glow turns out to be primarily about status or showing off income? More giving might mean greater status for one person, but less for another. In this sort of a zero-sum game, it might actually make sense to discourage charitable contributions and rely instead on higher taxes.

Still other questions involve the interactions between the motives of givers and the goals of the non-profit-making organizations that actually spend the money to increase the public good. Fundraisers often say that "people give because we ask them to give." What effect does the "ask" have on the utility of the giver? Given the lengths that some people will go to in order to avoid having to say no, it seems unlikely that the process that leads to charitable giving always involves increases in the utility of those who contribute. It seems that neuroeconomic techniques are particularly well suited to addressing these sorts of questions.

Broadening the perspective a little, neural-level results such as the ones discussed in this chapter might provide important hints regarding how people develop the motivational tendencies that apply in interpersonal and public good exchanges. Particularly revealing is the finding that activity in midbrain areas can accompany altruistic behavior. After all, in terms of learning, the primary function of a reward is to strengthen behavioral patterns that lead to the reward and weaken those that do not. Altruistic behavior may well be subject to the same type of learning, as are other types of behavior. So, an important question is, to what degree is there an *a priori* link between altruistic behavior and reward, or is engagement in altruistic behavior more an acquired taste? For example, an interesting hypothesis is that a warm-glow related reward can serve as a primary learning reinforcer, which over time generalizes to provide pure altruistic rewards. Praising toddlers for sharing their toys might deliver the warm-glow benefit, which then promotes future altruistic behavior even when no one is looking.

References

Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behav. Cogn. Neurosci. Rev.* 1, 21.

Amodio, D.M. and Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277.

Andreoni, J. (1989). Giving with impure altruism: applications to charity and Ricardian equivalence. *J. Political Econ.* 97, 1447–1458.

Andreoni, J. (1990). Impure altruism and donations to public good: a theory of warm-glow giving. *Economic J.* 100, 464–477.

Andreoni, J. (1995). Cooperation in public good experiments: kindness or confusion. *Am. Econ. Rev.* 85, 891–904.

Andreoni, J. (2006). Philanthropy. *Handbook of the Economics of Giving, Reciprocity and Altruism*. Amsterdam: Elsevier, pp. 1201–1269.

Andreoni, J., Harbaugh, W., and Vesterlund, L. (2003). The carrot or the stick: rewards, punishments, and cooperation. *Am. Econ. Rev.* 93, 893–902.

Andreoni, J., Harbaugh, W.T., and Vesterlund, L. (2008). Altruism in experiments. In: S.N. Durlauf and L.E. Blume (eds), *The New Palgrave Encyclopedia of Economics*, 2nd edn. Basingstoke: Palgrave Macmillan, (available online at http://www.dictionaryofeconomics.com/article?id=pde2008_A000240 doi:10.1057/9780230226203.0035)..

Ariely, D., Bracha, A., and Meier, S. (2007). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. IZA Discussion Paper No. 2968 (available online at SSRN: http://ssrn.com/abstract=1010620).

Axelrod, R.M. (1984). *The Evolution of Cooperation*. New York, NY: Basic Books.

Bateson, M., Nettle, D., and Roberts, G. (2006). Cues of being watched enhance cooperation in real-world setting. *Biol. Letts* 2, 412–414.

Batson, C.D. (1991). Evidence for altruism: toward a pluralism of prosocial motives. *Psychological Inquiry* 2, 107–122.

Batson, C.D., Dyck, J.L., Brandt, J.R. *et al.* (1988). Five studies testing two new egoistic alternatives to the empathy-altruism hypothesis. *J. Pers. Social Psychol.* 55, 52–77.

Becker, G.S. (1974). A theory of social interactions. *J. Political Econ.* 82, 1063–1093.

Bekkers, R. and Wiepking, P. (2007). Generosity and Philanthropy A Literature Review. 2007. Available online at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1015507

Bering, J.M., McLeod, K., and Shackelford, T.K. (2005). Reasoning about dead agents reveals possible adaptive trends. *Hum. Nature* 16, 360–381.

Bergstrom, T., Blume, L., and Varian, H. (1986). On the private provision of public goods. *J. Public Econ.* 29, 25–49.

Blair, R.J. (2005). Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. *Conscious Cogn.* 14, 698–718.

Carstensen, L.L. (2006). The influence of a sense of time on human development. *Science* 30, 1913–1915.

Charney, D.S. (2004). Psychobiological mechanisms of resilience and vulnerability: implications for successful adaptation to extreme stress. *Am. J. Psych.* 161, 216.

de Quervain, D.J.F., Fischbacher, U., Treyer, V. *et al.* (2004). The neural basis of altruistic punishment. *Science* 305, 1254–1258.

Delgado, M.R., Frank, R.H., and Phelps, E.A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8, 1611–1618.

Depue, R.A. and Collins, P.F. (1999). Neurobiology of the structure of personality: dopamine, facilitation of incentive motivation, and extraversion. *Behav. Brain Sci.* 22, 491–517.

Depue, R.A. and Morrone-Strupinsky, J.V. (2005). A neurobehavioral model of affiliative bonding: Implications for conceptualizing a human trait of affiliation. *Behav. Brain Sci.* 28, 313–350.

Duval, S. and Wicklund, R.A. (1972). *A Theory of Objective Self Awareness*. New York, NY: Academic Press.

Eckel C. , and Grossman P. (2004). Giving to Secular Causes by the Religious and Nonreligious: An Experimental Test of the

Responsiveness of Giving to Subsidies. *Nonprofit and Voluntary Sector Quarterly* 33: 271-89.

Glazer, A. and Konrad, K.A. (1996). A signaling explanation for charity. *Am. Econ. Rev.* 86, 1019–1028.

Goeree, J.K., Holt, C.A., and Laury, S.K. (2002). Private costs and public benefits: unraveling the effects of altruism and noisy behavior. *J. Public Econ.* 83, 255–276.

Greene, J.D., Sommerville, R.B., Nystrom, L.E. *et al*. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science* 14, 2105–2108.

Hafer, C.L. and Begue, L. (2005). Experimental research on just-world theory: problems, developments, and future challenges. *Psychological Bull.* 131, 128–167.

Harbaugh, W.T. (1998a). The prestige motive for making charitable transfers. *Am. Econ. Rev.* 88, 277–282.

Harbaugh, W.T. (1998b). What do donations buy? *J. Public Econ.* 67, 269–284.

Harbaugh, W.T., Mayr, U., and Burghart, D.R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* 316, 1622.

Isaac, R.M. and Walker, J.M. (1988). Group size effects in public goods provision: the voluntary contributions mechanism. *Q. J. Economics* 103, 179–199.

King-Casas, B., Tomlin, D., Anen, C. *et al*. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83.

Kingma, B.R. and McClelland, R. (1995). Public radio stations are really, really not public goods: charitable contributions and impure altruism. *Ann. Public Coop. Econ.* 66, 65–76.

Knafo, A., Israel, S., Darvasi, A. *et al*. (2007). Individual differences in allocation of funds in the dictator game associated with length of the arginine vasopressin 1a receptor RS3 promoter region and correlation between RS3 length and hippocampal mRNA. Available online at doi:10.1111/j.1601-183X.2007.00341.x.

Knoch, D., Pascual-Leone, A., Meyer, K. *et al*. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314, 829.

Knutson, B. and Cooper, J.C. (2005). Functional magnetic resonance imaging of reward prediction. *Curr. Opin. Neurol.* 18, 411–417.

Koenigs, M., Young, L., Adolphs, R. *et al*. (2007). Damage to the prefrontal cortex increases utilitarian moral judgments. *Nature* 446, 908–911.

Kolm, S.C. (1969). *The Optimal Production of Justice. Public Economics*. London: Macmillan.

Lansdell, H. (1906). *The Sacred Tenth or Studies in Tithe-Giving Ancient and Modern*. Society for Promoting Christian Knowledge.

Larkin, G.R.S., Gibbs, S.E.B., Khanna, K. *et al*. (2007). Anticipation of monetary gain but not loss in healthy older adults. *Nat. Neurosci.* 10, 787–791.

Lerner, M. (1980). *The Belief in a Just World: A Fundamental Delusion*. New York, NY: Plenum Press.

Levin, P.F. and Isen, A.M. (1975). Further studies on the effect of feeling good on helping. *Sociometry* 38, 141–147.

Moll, J., Krueger, F., Zahn, R. *et al*. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *PNAS* 103, 15,623.

Morewedge, C.K. and Schooler, J. (2008). Mind diffusion: when attribution of mind is greater to individuals than to group members. Working Paper, Carnegie Mellon University.

Norenzayan, A. and Shariff, A. (2007). God is watching you: priming God concepts increases prosocial behavior in an anonymous economic game. *Psychological Sci.* 18, 803–809.

O'Doherty, J.P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr Opin Neurobiol* 14, 769–776.

Palfrey, T.R. and Prisbrey, J.E. (1997). Anomalous behavior in public goods experiments: how much and why. *Am. Econ. Rev.* 87, 829–846.

Raz, N. (2005). The aging brain observed in vivo: differential changes and their modifiers. In: R. Cabeza, L. Nyberg, and D. Park (eds), *Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral Aging*. New York, NY: Oxford University Press, pp. 19–57.

Ribar, D.C. and Wilhelm, M.O. (2002). Altruistic and joy-of-giving motivations in charitable behavior. *J. Political Econ.* 110, 425–457.

Ruby, P. and Decety, J. (2004). How would you feel versus how do you think she would feel? A neuroimaging study of perspective-taking with social emotions. *J. Cogn. Neurosci.* 16, 988–999.

Samuelson, P.A. (1954). The pure theory of public expenditure. *Rev. Econ. Stats* 36, 387–389.

Silvia, P.J. and Duval, T.S. (2001). Objective self-awareness theory: recent progress and enduring problems. *Pers. Social Psychol. Rev.* 5, 230.

Singer, T., Seymour, B., O'Doherty, J. *et al*. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science* 303, 1157–1162.

Slovic, P. (2007). If I look at the mass I will never act. Psychic numbing and genocide. *Judgment Decision-Making* 2, 79–95.

Sosis, R. and Alcorta, C. (2003). Signaling, solidarity, and the sacred: the evolution of religious behavior. *Evol. Anthropol.* 12, 264–274.

Tankersley, D., Stowe, C.J., and Huettel, S.A. (2007). Altruism is associated with an increased neural response to agency. *Nat. Neurosci.* 10, 150–151.

Wilson, D.S. (2002). *Darwin's Cathedral*. Chicago, IL: University of Chicago Press.

Zak, P.J., Stanton, A.A., and Ahmadi, S. (2007). Oxytocin increases generosity in humans. *PLoS ONE* 2(11), e1128, doi:10.1371/journal.pone.0001128.

Zysset, S., Huber, O., Ferstl, E., and von Cramon, D.Y. (2002). The anterior frontomedian cortex and evaluative judgment: an fMRI study. *NeuroImage* 15, 983–991.

PART IV

UNDERSTANDING VALUATION – LEARNING VALUATIONS

CHAPTER

21

Midbrain Dopamine Neurons: A Retina of the Reward System?

Wolfram Schultz

OUTLINE

INTRODUCTION

We can define rewards as objects or events that generate approach and consummatory behavior, produce learning of such behavior, represent positive outcomes of economic decisions, and may engage positive emotions. Rewards are crucial for individual and gene survival and support elementary processes, such as drinking, eating, and reproduction. This behavioral definition attributes reward function also to certain non-alimentary and non-sexual entities, including money, technical artifacts, esthetic stimulus attributes, and mental events. Rewards engage agents in such diverse behaviors as foraging, and trading on stock markets.

The functions of primary sensory systems are based on the physics and chemistry of specific stimuli impinging on their receptors. By contrast, rewards are often polysensory and do not engage specialized receptors. Reward functions are not defined primarily by the physics and chemistry of their objects, but by the behavioral reactions they induce. In order to use rewards for specific behavioral reactions, the brain needs to extract the reward information from heterogeneous mechanical, electromagnetic, chemical, and other events impinging on our sensory organs.

In the absence of dedicated reward receptors, it would be helpful for the brain to have an explicit neuronal signal that identifies a reward object and contains information about its motivational or economic value, irrespective of other sensory properties. Such a reward signal could provide an input to neuronal processes underlying approach behavior, learning, and economic decision making, just as, for example, visual responses in retinal ganglion cells identify specific electromagnetic events and deliver this information to brain centers involved in visual perception and sensorimotor reactions.

This chapter describes how the search for a "retina of the reward system" has revealed phasic signals in midbrain dopamine neurons that reflect information about the motivational or economic value of rewards regardless of other stimulus components and irrespective of specific behavioral reactions necessary to obtain the rewards.

CONCEPTS OF REWARD

In the absence of dedicated reward receptors, the behavioral definition of reward allows us to use behavioral theories for investigating the neural processing of reward. In analogy to David Marr's suggestion for the visual system (Marr, 1982), behavioral theories help us to define and understand the problem to be solved (approach behavior, learning, and decision making), and neurobiology helps us to understand how the problem is being solved (neural reward mechanisms).

Reward Value, Utility and Risk

Rewards are positive outcomes that have specific magnitudes and occur with specific probabilities. Agents make choices between options whose motivational or economic values are determined by the magnitudes and probabilities of outcomes, all other outcome parameters being equal.

The term "mean (of reward)" seems more appropriate than "expectation (of reward)". Monkeys may not have a notion of mathematical expectation, and instead rely on experienced population or sample means. In addition the term "expectation" is used to denote a psychological process. The term "risk" refers to a form of uncertainty in which the probability distributions are known, whereas "ambiguity" indicates uncertainty with incomplete knowledge of probabilities and is often referred to simply as "uncertainty." Risk in economics refers to the chances of losing and winning, rather than the narrower common sense meaning of the potential for loss.

Standard expected utility theory assumes that decision makers assign subjective utilities to objective outcome values, as measured objectively in choice preferences (Bernoulli, 1738). Preference is expressed as the probability of choosing one option over all others. Utility usually increases non-linearly as a function of value, and utility functions flatten gradually when utility increases progressively less with higher outcome values. Then, the losses from probabilistically occurring rewards appear larger than the gains, the mathematical expectation of the probability distribution of utilities (expected utility) drops with increasing variance, and the decision maker shows risk aversion. By contrast, assigning progressively more utility to higher outcome values favors variable outcomes and produces risk-seeking. Thus, the concave or convex shape of utility functions determines how much risk-sensitive individuals lose or gain in expected utility with risky outcomes. Only individuals with linear utility functions are risk-neutral. For more details, Chapter 3 of this volume.

Predictions

Evolutionary pressure demands the energy-efficient processing of information. One potential solution is to store predictions about future events at higher brain centers, process the differences between the newly incoming information and the predictions (prediction errors) in lower brain centers, and update the predictions in higher brain centers by forwarding only the less information-containing prediction errors rather than the full information about the state of the external world (Rao and Ballard, 1999). In this way, higher brain centers have access to full information about the external world for perceptions, decisions, and behavioral reactions at a much lower cost. This fundamental role of predictions would lead to two processes, namely learning, as defined by changes in behavior derived from updated predictions, and informed decision making through advance information about the available choice options (as opposed to guesses, when outcomes are unknown). As choice options are best described by probability distributions rather than singular values, reward predictions also comprise distributions, and specify the key parameters of mean or expected value, expected utility, and the corresponding SDs or variances.

Prediction Errors and Learning

"Prediction error" refers to the difference between the actual outcome and its prediction. Animal learning theory and efficient temporal difference (TD) reinforcement models postulate a crucial role for outcome prediction errors in Pavlovian and operant conditioning (Rescorla and Wagner, 1972; Sutton and Barto, 1981). Current views attribute Pavlovian learning to any acquisition of predictions, irrespective of vegetative reactions or striated muscle contractions, as long as the outcome is not conditional on a behavioral reaction. The crucial role of prediction errors is based on Kamin's blocking effect (1969), which demonstrates that learning and extinction advance only to the extent to which a reinforcer is better or worse than predicted, and slow as the prediction approaches the reinforcer.

DOPAMINE REWARD-RELATED RESPONSES

Pure Reward Signals

Pure reward neurons would carry a signal with abstract information about the objective economic reward value or utility that is useful for eliciting the known behavioral effects of rewards (approach behavior, learning, decision making, and positive emotion). Such a pure reward signal should not be specific for the sensory properties of the reward, and thus not distinguish between individual reward objects; nor should it reflect the specifics of the behavioral reaction necessary to obtain the reward. However, some reward-related neurons do carry such additional information – for example, action-value neurons that combine specific movement information with specific reward information.

Dopamine Response to Reward

The majority of midbrain dopamine neurons (75–80%) show rather homogeneous, phasic activations with latencies of <100 ms and durations of <200 ms to temporally unpredicted food and liquid rewards. The remaining neurons are unresponsive to all stimuli tested. The responses increase monotonically with reward magnitude, such as liquid volume (Tobler *et al*. 2005), apparently coding reward value or utility. Although further tests are necessary, the dopamine reward responses seem to constitute pure reward signals.

Prediction-error Coding

The dopamine response to the reward appears to code a prediction error, such that a reward that is better than predicted elicits an activation (positive prediction error), a fully predicted reward draws no response, and a reward that is worse than predicted induces a depression (negative error). Thus, the dopamine response implements fully the crucial term of the Rescorla-Wagner learning model.

The error response varies quantitatively with the mean reward, irrespective of specific combinations of reward magnitude and probability (Fiorillo *et al*., 2003; Satoh *et al*., 2003; Morris *et al*., 2004; Nakahara *et al*., 2004; Bayer and Glimcher, 2005; Tobler *et al*., 2005). The prediction-error response is sensitive to the time of the reward, as a delayed reward induces a depression at its original time and an activation at its new time. Whereas the error coding is evident for activations reflecting positive prediction errors, it is more difficult to assess for the negative-error related depressions with their narrow dynamic range unless the full period of depression is considered (Bayer *et al*., 2007).

Stringent Tests for Prediction Error

The coding of reward prediction errors should be tested in stringent paradigms developed by animal learning theory. In the blocking test (Kamin, 1969), a stimulus that is paired with a fully predicted reward does not become a valid reward predictor. The absence of a reward following the blocked stimulus does not produce a prediction error, or a response in dopamine neurons, even after extensive stimulus–reward pairings (Waelti *et al*., 2001). By contrast, the delivery of a reward after a blocked stimulus constitutes a positive prediction error and, accordingly, elicits a dopamine activation.

In the conditioned inhibition paradigm, a test stimulus is presented simultaneously with an established reward-predicting stimulus, but no reward is given after the compound, making the test stimulus a predictor for the absence of reward. Reward omission after such a conditioned inhibitor does not produce a negative prediction error or a depression in dopamine neurons (Tobler *et al*., 2003). By contrast, delivery of a reward after the inhibitor produces a strong positive prediction error and, accordingly, a strong dopamine activation.

The results from these formal paradigms confirm that dopamine neurons indeed show bidirectional coding of reward-prediction errors.

Adaptive Coding

The dopamine response does not simply reflect the prediction error of the mean reward, but also adapts to the predicted reward distribution within 2 seconds. When visual stimuli predict specific binary distributions of equiprobable reward magnitudes ($P = 0.5$ each), delivery of reward with the larger of the two magnitudes within each distribution elicits the same dopamine activation, despite 10-fold differences in reward magnitude and prediction error (Tobler *et al.*, 2005). Experimental separation of mean and standard deviation (SD) suggests that the dopamine response codes the prediction error divided by, or normalized to, the SD, reflecting how much the delivered reward differs from its prediction in units of SD rather than mean. Theoretical considerations suggest that error teaching signals that are scaled by risk rather than mean can mediate stability in learning that is resistant to the predicted risk of outcomes (Preuschoff and Bossaerts, 2007; see also Chapter 23 of this volume).

Adaptive shifts in responses to predicted outcome distributions also occur in the orbitofrontal cortex and striatum of monkeys (Tremblay and Schultz, 1999; Cromwell *et al.*, 2005; Hosokawa et al., 2007) and humans (Breiter *et al.*, 2001; Nieuwenhuis *et al.*, 2005). However, shifts are not observed in all orbitofrontal reward neurons, nor do they occur when there is insufficient time for adaptation (Padoa-Schioppa and Assad, 2008). Future experiments might search for a potential common mechanism across reward structures that may match the probability distributions of neuronal responses to the probability distributions of rewards.

Dopamine Response to Reward-predicting Stimuli

Dopamine neurons are activated by reward-predicting visual, auditory, and somatosensory stimuli (for review, see Schultz, 1998). These responses occur irrespectively of the sensory modalities and spatial positions of stimuli, and of arm, mouth, and eye movements being executed. The responses increase with the means of predicted reward probability distributions, without distinguishing between magnitude and probability (Tobler *et al.*, 2005). Response magnitude increases with decreasing reaction time, and thus may incorporate the animal's motivation (Satoh *et al.*, 2003). In choices between different reward values or delays, the responses to the presentation of choice options reflect the animal's future chosen reward (Morris *et al.*, 2006) or the highest reward across options (Roesch *et al.*, 2007).

During the course of learning, the dopamine response to the reward decreases gradually, and a response to the reward-predicting stimulus develops (Mirenowicz and Schultz, unpublished; Takikawa *et al.*, 2004; Pan *et al.*, 2005). The acquisition of conditioned responding is sensitive to blocking, indicating that prediction errors play a role in the learning of dopamine neurons. The response transfer to reward-predicting stimuli complies with the principal characteristics of teaching signals of efficient reinforcement models (Sutton and Barto, 1981). The response shift does not involve the back-propagation of prediction errors of earlier TD models (Montague *et al.*, 1996), and is reproduced in recent TD implementations (Suri and Schultz, 1999; Pan *et al.*, 2005).

Physically intense stimuli with attention-inducing and possibly rewarding functions induce activations in dopamine neurons. These responses are enhanced by stimulus novelty (Ljungberg *et al.*, 1992; Horvitz *et al.*, 1997), but disappear quickly with aversive conditioning (Mirenowicz and Schultz, unpublished). However, attention-inducing reward omission induces only depressions. Other strong attentional stimuli, such as punishers and conditioned inhibitors, induce predominantly depressions and only brief, non-differential, initial activation components, and rarely full neuronal activations (Mirenowicz and Schultz, 1996; Schultz *et al.*, 1997; Tobler *et al.*, 2003). Some of these non-reward-related activations may be due to response generalization. Responsiveness to non-reward-predicting stimuli increases in contexts with higher proportions of stimuli predicting reward as opposed to non-reward (Schultz and Romo, 1990; Waelti *et al.*, 2001; Tobler *et al.*, 2003).

Taken together, these characteristics may qualify the main dopamine activations following reward-predicting stimuli as pure reward signals, with the possible exception of an additional motivational component. The responses to non-reward-predicting stimuli also occur irrespectively of sensory modalities, including responses to intense stimuli, initial non-differential response components to reward-predicting stimuli, and generalized responses; their functions may be related to the positively motivating attributes of intense, novel, and reward-resembling stimuli, but do not seem to reflect general arousing or attentional functions irrespective of motivational valence.

Dopamine Concentrations

Fast-scan cyclic voltammetry detects changes in dopamine concentration in dopamine projection areas downsteam of impulse activity of dopamine cell bodies, at time-courses somewhat comparable with electrophysiology. Dopamine concentrations increase after unpredicted food rewards, shift to reward-predicting stimuli during conditioning, and show a propensity for generalization to non-reward-predicting stimuli in ventral striatum (Day *et al.*, 2007). These data replicate some of the essential reward responses seen with electrophysiology, and suggest that the impulse responses indeed lead to dopamine release.

Reward Signals in Other Brain Structures

Among rather heterogeneous neuronal populations, a fraction of neurons in orbitofrontal cortex and striatum are activated in relation to rewards and reward-predicting stimuli (Thorpe *et al.*, 1983; Ravel *et al.*, 2003). A subset of orbitofrontal reward neurons show graded responses with reward magnitude (Wallis and Miller, 2003), and some of them carry pure reward signals by reflecting the economic value of reward objects determined by choice preferences, without distinguishing between different rewards ("chosen value" neurons of Padoa-Schioppa and Assad, 2006). Most orbitofrontal and some striatal reward neurons do not discriminate between spatial positions or visual features of reward-related stimuli (Tremblay and Schultz, 1999; Hassani *et al.*, 2001).

However, other orbitofrontal and striatal reward neurons do not seem to carry pure reward signals, as they distinguish between different reward objects with same economic values ("offer value" neurons of Padoa-Schioppa and Assad, 2006) or are in addition sensitive to visual object features. Some neurons encode both reward and movement parameters, and are often viewed as parts of goal-directed mechanisms by which reward information labels specific movements with specific economic value, and thus may determine decisions between differently valued actions. Such neurons are found in the striatum (Hollerman *et al.*, 1998; Kawagoe *et al.*, 1998; Samejima *et al.*, 2005), and constitute the typical reward neurons found in prefrontal and parietal cortex (Watanabe, 1996; Platt and Glimcher, 1999; Sugrue *et al.*, 2004).

DOPAMINE REWARD RISK SIGNAL

If a reward signal reflects the mean reward scaled by the standard deviation (SD) of reward probability distributions, could there be a direct neuronal signal for risk as measured by SD? When varying reward probabilities from 0 to 1 while keeping magnitude constant, the mean increases monotonically with probability, whereas risk (SD, variance) and entropy show an inverted U function peaking at $P = 0.5$. More than one-third of dopamine neurons show a relatively slow, moderate activation that increases gradually in time between the reward-predicting stimulus and the reward, and varies monotonically with risk (Fiorillo *et al.*, 2003). The activation occurs in individual trials, and does not seem to constitute a prediction-error response propagating back from reward to the reward-predicting stimulus. The activation also increases monotonically with SD or variance when binary distributions of different equiprobable, non-zero reward magnitudes are used, entropy being constant at 1 bit. Thus, SD and variance appear to be viable measures for risk as coded by dopamine neurons. Risk-related activations have slower time-courses and lower peaks compared with the responses to reward-predicting stimuli and prediction errors at different behavioral periods.

A dopamine risk signal may have three functions. First, it might influence the scaling of responses by SD in dopamine and other reward neurons. Second, it might provide an input to brain structures dealing with the assessment of reward risk *per se*. Third, it might combine with a mean economic value signal to represent information about the expected utility in risk-sensitive individuals, according to the mean-variance concept in financial decision theory (Levy and Markowitz, 1958).

DOPAMINE CONFUSIONS

Neurotransmitters can influence postsynaptic neurons through phasic signals transmitting time-specific information for specific behavioral processes, and through tonic, permissive actions, without much variation over time. Dopamine seems to have both actions (for review, see Schultz, 2007).

The phasic, informational signals of dopamine neurons relate to reward and risk, and are compatible with subsecond time-courses of rapid behavioral processes. They produce subsecond rises of dopamine concentrations well above 100 nM. These signals are suitable for assigning reward credit to stimuli and actions during learning by enabling synaptic modifications, and they may inform rapid neuronal decision mechanisms. At mostly slower timescales of seconds, a negative dopamine signal reports aversive events, and voltammetrically measured dopamine is released with movements. Microdialysis reveals moderate dopamine fluctuations over several minutes with punishment and stress, but these are much slower than in most learning and decision-making mechanisms.

In addition to its informational function, dopamine appears to play a permissive role devoid of specific, behavioral information that is required for a large variety of motor, cognitive, motivational, and learning processes to function properly. These processes become deficient when dopamine transmission is impaired by lesions or receptor antagonists, and recover partly after pharmacological stimulation of dopamine receptors without the phasic informational signals being restored. This role involves sustained, finely controlled, and only slowly varying extracellular dopamine concentrations in the order of 5–20 nM that exert a necessary, sustained influence on postsynaptic striatal and cortical receptors, membranes, and secondary messenger systems processing a large variety of specific, informational, behavior-related signals. The permissive role would allow the limited number of dopamine neurons to be involved in a large variety of functions residing in postsynaptic structures like striatum and cortex, without providing them with phasic input signals containing specific information.

Confusion will arise when confounding the different informational and permissive modes of dopamine action and trying to identify a single behavioral dopamine function to the exclusion of other functions, as we usually do for primary sensory systems. For example, strong dopamine reductions in the striatum lead to Parkinsonian movement deficits, but dopamine neurons do not encode movement parameters. Apparently, dopamine has a permissive effect on movement processes residing in the striatum without carrying its own movement-related information, and Parkinsonism may reflect a disorder of dopamine-dependent, movement-related activity in striatal neurons rather than a deficit in informational coding of dopamine neurons. Likewise, impaired inputs from dopamine neurons to the ventral striatum induce deficiencies in incentive motivational salience, but dopamine neurons do not carry a phasic incentive salience signal, as they are activated very little by punishers, which are motivationally very salient. Apparently, the deficits in incentive salience are due to a disfunctional ventral striatum following impairment of the permissive rather than informational dopamine action. Thus, dopamine is involved in movement and incentive salience, but this involvement reflects the permissive rather than the informational reward- and risk-related dopamine function. Postsynaptic mechanisms would be able to discriminate between the permissive and informational dopamine functions on the basis of their time-courses and instantaneous dopamine concentrations.

CONCLUSIONS

Just as the retina transfers information from electromagnetic waves to action potentials, dopamine neurons transfer information from polysensory stimuli to action potentials signaling rewards. The reward signal informs neural decision-making mechanisms about the abstract motivational or economic reward value, or possibly utility, as the key decision variable for economic choices. Neurons in other reward structures carry additional reward information, including reference to drive states and sensory details about individual reward objects. However, the economic reward value signal of dopamine neurons codes neither simply reward nor mean economic value. Rather, it encodes reward outcome as it deviates from prediction, thus fulfilling the basic requirement for a prediction-error teaching signal postulated by learning theory, and it is scaled in units of standard deviation rather than mean. A separate, slower dopamine signal informs directly about the degree of risk. Whereas the dopamine reward and risk signals are fast enough for immediate behavioral reactions, dopamine also operates with slower time-courses in a much wider spectrum of behavior, often in a permissive role without much temporal and behavioral information.

Acknowledgments

Our work was supported by the Wellcome Trust, the Swiss National Science Foundation, the Human Frontiers Science Program, and several other grant and fellowship agencies.

References

Bayer, H.M. and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141.

Bayer, H.M., Lau, B., and Glimcher, P.W. (2007). Statistics of dopamine neuron spike trains in the awake primate. *J. Neurophysiol.* 98, 1428–1439.

Bernoulli, D. (1738). Specimen theoriae novae de mensura sortis. Comentarii Academiae Scientiarum Imperialis Petropolitanae. *Papers Imp. Acad. Sci. St Petersburg* 5, 175–192, Translated as: (1954) Exposition of a new theory on the measurement of risk. *Econometrica* 22, 23–36.

Breiter, H.C., Aharon, I., Kahneman, D. *et al*. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30, 619–639.

Cromwell, H.C., Hassani, O.K., and Schultz, W. (2005). Relative reward processing in primate striatum. *Exp. Brain Res.* 162, 520–525.

Day, J.J., Roitman, M.F., Wightman, R.M., and Carelli, R.M. (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.* 10, 1020–1028.

Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902.

Hassani, O.K., Cromwell, H.C., and Schultz, W. (2001). Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J. Neurophysiol.* 85, 2477–2489.

Hollerman, J.R., Tremblay, L., and Schultz, W. (1998). Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J. Neurophysiol.* 80, 947–963.

Horvitz, J.C., Stewart, T., and Jacobs, B.L. (1997). Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. *Brain Res.* 759, 251–258.

Hosokawa, T., Kato, K., Inoue, M., and Mikami, A. (2007). Neurons in the macaque orbitofrontal cortex code relative preference of both rewarding and aversive outcomes. *Neurosci. Res.* 57, 434–445.

Kamin, L.J. (1969). Selective association and conditioning. In: N.J. Mackintosh and W.K. Honig (eds), *Fundamental Issues in Instrumental Learning*. Dalhousie, NB: Dalhousie University Press, pp. 42–64.

Kawagoe, R., Takikawa, Y., and Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* 1, 411–416.

Levy, H. and Markowitz, H.M. (1979). Approximating expected utility by a function of mean and variance. *Am. Econ. Rev.* 69, 308–317.

Ljungberg, T., Apicella, P., and Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* 67, 145–163.

Marr, D. (1982). *Vision*. San Francisco, CA: Freeman.

Mirenowicz, J. and Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379, 449–451.

Montague, P.R., Dayan, P., and Sejnowski, T.J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.

Morris, G., Arkadir, D., Nevet, A. *et al.* (2004). Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* 43, 133–143.

Morris, G., Nevet, A., Arkadir, D. *et al.* (2006). Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* 9, 1057–1063.

Nakahara, H., Itoh, H., Kawagoe, R. *et al.* (2004). Dopamine neurons can represent context-dependent prediction error. *Neuron* 41, 269–280.

Nieuwenhuis, S., Heslenfeld, D.J., Alting van Geusau, N. *et al.* (2005). Activity in human reward-sensitive brain areas is strongly context dependent. *NeuroImage* 25, 1302–1309.

Padoa-Schioppa, C. and Assad, J.A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.

Padoa-Schioppa, C. and Assad, J.A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat. Neurosci.* 11, 95–102.

Pan, W.-X., Schmidt, R., Wickens, J.R., and Hyland, B.I. (2005). Dopamine cells respond to predicted events during classical conditioning: Evidence for eligibility traces in the reward-learning network. *J. Neurosci.* 25, 6235–6242.

Platt, M.L. and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.

Preuschoff, K. and Bossaerts, P. (2007). Adding prediction risk to the theory of reward learning. *Ann. NY Acad. Sci.* 1104, 135–146.

Rao, R.P.N. and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87.

Ravel, S., Legallet, E., and Apicella, P. (2003). Responses of tonically active neurons in the monkey striatum discriminate between motivationally opposing stimuli. *J. Neurosci.* 23, 8489–8497.

Rescorla, R.A. and Wagner, A.R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: A.H. Black and W.F. Prokasy (eds), *Classical Conditioning II: Current Research and Theory*. New York, NY: Appleton Century Crofts, pp. 64–99.

Roesch, M.R., Calu, D.J., and Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* 10, 1615–1624.

Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340.

Satoh, T., Nakai, S., Sato, T., and Kimura, M. (2003). Correlated coding of motivation and outcome of decision by dopamine neurons. *J. Neurosci.* 23, 9913–9923.

Schultz, W. (2007). Multiple dopamine functions at different time courses. *Ann. Rev. Neurosci.* 30, 259–288.

Schultz, W. and Romo, R. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioral reactions. *J. Neurophysiol.* 63, 607–624.

Schultz, W., Dayan, P., and Montague, R.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.

Suri, R. and Schultz, W. (1999). A neural network with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience* 91, 871–890.

Sutton, R.S. and Barto, A.G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychol. Rev.* 88, 135–170.

Takikawa, Y., Kawagoe, R., and Hikosaka, O. (2004). A possible role of midbrain dopamine neurons in short- and long-term adaptation of saccades to position-reward maping. *J. Neurophysiol.* 92, 2520–2529.

Thorpe, S.J., Rolls, E.T., and Maddison, S. (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp. Brain Res.* 49, 93–115.

Tobler, P.N., Dickinson, A., and Schultz, W. (2003). Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J. Neurosci.* 23, 10402–10410.

Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642–1645.

Tremblay, L. and Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature* 398, 704–708.

Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412, 43–48.

Wallis, J.D. and Miller, E.K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur. J. Neurosci.* 18, 2069–2081.

Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature* 382, 629–632.

CHAPTER

22

Theoretical and Empirical Studies of Learning

Yael Niv and P. Read Montague

OUTLINE

INTRODUCTION

One fundamental goal of behavioral neuroscience is to understand the decision making processes that animals and humans use in order to select actions in the face of reward and punishment. Recent research has pursued this goal from a variety of perspectives. In behavioral psychology, this question has been investigated through the paradigms of Pavlovian (classical) and instrumental (operant) conditioning, and much evidence has accumulated regarding the learned associations that control simple behaviors. Simple conditioning paradigms also form the backbone of neurobiological approaches to learning, where investigators seek underlying neural mechanisms. From a computational perspective, Pavlovian (passive) conditioning is modeled as *prediction learning* – learning the predictive (and sometimes causal) relationships between different events in the environment, such as the fact that the smell of food usually predicts that a tasty meal is forthcoming. Instrumental conditioning, on the other hand, involves learning to select actions that will bring about rewarding events and avoid aversive events. Computationally, such decision making is treated as attempting to *optimize* the consequences of actions. Thus, from an economic perspective, the study of instrumental conditioning is an inquiry into perhaps the most fundamental form of rational decision making.

Computational accounts of conditioned behavior have, in recent years, drawn heavily from a class of models called *reinforcement learning* (RL) models. These models, now commonly used in neurobiology, psychology, and machine learning, all share in common the use of a *scalar reinforcement signal* to direct learning. This is categorically different from, and lies in between, learning from an explicit teaching signal (as in "supervised learning" models, common in artificial intelligence applications), and learning of input statistics without any supervisory signal (as in "unsupervised learning" models, for instance of early visual processing; Lewicki and Olshausen, 1999). Thus algorithms and theory have been developed specifically for the case of RL.

Importantly, reinforcement learning provides a *normative framework* within which conditioning can be analyzed. That is, it suggests a means by which optimal prediction and action selection can be achieved, and exposes explicitly the computations that must be realized in the service of these. In contrast to descriptive models that describe behavior as it is, normative models study behavior from the point of view of its hypothesized *function* – that is, they study behavior *as it should be* if it were to accomplish specific goals in an optimal way. The appeal of normative models derives from several sources. First, because throughout evolution animal behavior has been shaped and constrained by its influence on fitness, one reasonable starting point is to view a particular behavior as an optimal or near-optimal adaptation to some set of problems (Kacelnik, 1997). Treating behavior as optimal allows for the generation of computationally explicit hypotheses that are directly testable. Second, discrepancies between observed behavior and the predictions of normative models are often illuminating: these can shed light on the neural and/or informational constraints under which animals make decisions, or suggest that animals are, in fact, optimizing something other than what the model has assumed. Such approaches are familiar to economists. In economics, it is frequently assumed that a decision maker is rational (even though we know people are not), and deviations from rationality are used to devise progressively more accurate theories of human decision making.

Finally, as has been recently the case for economic constructs as well, the relevance of reinforcement learning models to human and animal decision making has been strengthened by measurements of neural correlates of some of the major constituents of these models. Specifically, extracellular recordings in behaving animals and functional imaging of human decision making have revealed in the brain the existence of a key reinforcement learning signal, the *temporal difference reward-prediction error*.

In this chapter, we introduce the formal reinforcement learning framework and give a brief background to the origins and history of reinforcement learning models of decision making (for a comprehensive textbook account of RL methods, see Sutton and Barto, 1998). In the second section, we review the multiple lines of evidence linking reinforcement learning to the function of dopaminergic neurons in the mammalian midbrain. These data demonstrate the strength of the computational model and normative framework for interpreting and predicting a wide range of (otherwise confusing) neural activity patterns. The third section extends these results to more recent data from human imaging experiments. In these experiments, the combination of reinforcement learning models of choice behavior and online imaging techniques has allowed researchers to detect in the brain the existence of specific "hidden variables" controlling behavior (such as the subjective value of different options). Prior to this work such variables could only be postulated, or worse, were presumed to have descriptive ("as if") status alone. Together, the results from the latter-two sections put on firmer ground concepts central to neuroeconomics – for example, expected utility in units of a common currency and temporal discounting. In the fourth section, we discuss aspects of learning not associated with phasic dopamine signals, such as learning about the vigor (or rate) with which actions should be performed (which seems to be related more to tonic levels of dopamine in the striatum), and adapting learning rates to the natural statistics of the learning problem (which has been associated with the functions of the neuromodulators acetylcholine and norepinephrine). The final section discusses some of the fundamental limitations of the current theories of learning, and highlights questions for future research.

REINFORCEMENT LEARNING: THEORETICAL AND HISTORICAL BACKGROUND

Historically, the core ideas in reinforcement learning arose from two separate and parallel lines of research. One axis is mainly associated with Richard Sutton, formerly a psychology major, and his PhD advisor, Andrew Barto, a computer scientist. Interested in artificial intelligence and agent-based learning, Sutton and Barto developed algorithms for reinforcement learning that were inspired by the psychological literature on Pavlovian and instrumental conditioning (Sutton, 1978; Barto *et al.*, 1983; Sutton and Barto, 1990, 1998). Stemming from a different background, the second axis was led by electrical engineers such as Dimitri Bertsekas and John Tsitsiklis. Within the fields of operations research and optimal control, Bertsekas and Tsitsiklis developed stochastic approximations to dynamic programming (which they termed "neurodynamic programming") that led to similar reinforcement learning rules (e.g. Bertsekas and Tsitsiklis, 1996). The fusion of these two lines of research reinterpreted the behaviorally-inspired and somewhat heuristically-derived reinforcement learning algorithms in terms of optimality, and provided tools for analyzing their convergence properties in different situations.

The Rescorla-Wagner Model

The early impetus for the artificial intelligence trajectory can be traced back to the behaviorist movement in psychology in the early twentieth century. Behaviorism played an important "rigorizing" role for the kinds of experiments and data that psychologists would come to find acceptable, but was largely a-theoretic. This gave rise to the field of "mathematical psychology" in the 1950s, within which statistical models of learning were considered for the first time. In a seminal paper that helped to establish this field, Bush and Mosteller (1951) developed one of the first detailed formal accounts of learning. Together with Kamin's (1969) idea that learning should occur only when outcomes are "surprising," the Bush and Mosteller "linear operator" model finds its most popular expression in the now-classic Rescorla-Wagner model of Pavlovian conditioning (Rescorla and Wagner, 1972). The Rescorla-Wagner model, arguably the most influential model of animal learning to date, explained the puzzling behavioral phenomena of blocking, overshadowing, and conditioned inhibition by postulating that learning occurs *only when events violate expectations*. For instance, in a conditioning trial in which *conditional stimuli* CS_1 and CS_2 (say, a light and a tone) were presented, as well as an affective stimulus such as food or a tail-pinch (termed the *unconditional stimulus*; US), Rescorla and Wagner postulated that the associative strength of each of the conditional stimuli $V(CS_i)$ will change according to

$$V_{new}(CS_i) = V_{old}(CS_i) + \eta(CS_i, US) \times \left[\lambda(US) - \sum_i V_{old}(CS_i)\right] \qquad (22.1)$$

In this *error-correcting* learning rule, learning is driven by the discrepancy between what was predicted ($\Sigma_i V(CS_i)$ where i indexes all the CSs present in the trial) and what actually happened ($\lambda(US)$, whose magnitude is related to the worth of the unconditional stimulus). $\lambda(US)$ quantifies the maximal associative strength that the unconditional stimulus can support and $\eta(CS_i, US)$ is a learning rate that can depend on the salience properties of both the conditional and the unconditional stimuli being associated.

From our perspective, the Rescorla-Wagner learning model was based on two the important (and innovative) assumptions or hypotheses: (1) learning happens only when events are not predicted, and (2) the predictions due to different conditional stimuli are summed together to form the total prediction in the trial. These assumptions allowed the model to explain parsimoniously several anomalous features of animal learning: why an already predicted unconditional stimulus will not support conditioning of an additional conditional stimulus (as in blocking; Kamin, 1969); why differently salient conditional stimuli presented together might become differentially associated with an unconditional stimulus (as in overshadowing; Reynolds, 1961); and why a stimulus that predicts the *absence* of an expected unconditional stimulus acquires a negative associative strength (as in inhibitory conditioning; Konorski, 1948; Rescorla and Lolordo, 1968).

The Rescorla-Wagner model explains quite elegantly a large collection of behavioral data (and, furthermore, predicted previously undetected phenomena such as over-expectation; Rescorla, 1970; Kremer, 1978); however, it suffers from two major shortcomings. First, by treating the conditional and unconditional stimuli as qualitatively different, it does not extend to the important phenomenon of second-order conditioning. Second-order conditioning is a behavioral analog to the idea of transitivity: if stimulus B predicts an affective outcome (say, reward) and stimulus A comes to predict stimulus B, then stimulus A also gains predictive value, i.e., *a predictor of a predictor is a predictor*. To be more concrete, suppose that a tone is repeatedly followed by food delivery so that an association forms between them (tone predicts food), then subsequent pairing of a light with the tone can confer predictive value to the light (for instance, the animal will approach the light and perhaps salivate when it comes on) – this effect is second-order conditioning. This laboratory paradigm is especially important given the prevalence of conditioning of humans to monetary outcomes, which are second-order predictors of a wide range of affectively desirable unconditional stimuli, such as food and shelter.

The second shortcoming of the Rescorla-Wagner rule is that its basic unit of learning is (the artificial contrivance of) a conditioning *trial* as a discrete temporal object. Not only does this impose an experimenter-oriented parsing of otherwise continuous events, but it also fails to account for the sensitivity of conditioning to the precise temporal relations between the conditional and the unconditional stimuli within a trial (that is, whether they appeared simultaneously or serially, their order of appearance, and whether there was a time lag between them).

Temporal Difference Learning

To overcome these two problems, Sutton and Barto (1990) suggested the *temporal-difference learning rule* as a model of Pavlovian conditioning (i.e., prediction learning). Temporal-difference (TD) learning is quite simply a temporally extended version of the Rescorla-Wagner model discussed above. However, although the distinctions between the Rescorla-Wagner model and the TD model will seem subtle, the differences allow the TD model to account for higher-order conditioning and make it sensitive to the (potentially) important temporal relationships within a learning trial (Sutton and Barto, 1990).

In TD learning, the goal of the learning system (agent) is to estimate the future value of different states. For example, from a learning standpoint, the TD model assumes that the goal of a rat running about in a novel arena containing hidden rewards (e.g. food pellets hidden underneath bedding) is to learn the value of various positions in the arena. One way to evaluate the locations would be to estimate for each location the total amount of reward that the rat could expect to receive in the distant future. However, after a location is visited, there are many paths leading away from it, which yield variable amounts of reward. According to TD learning, a useful value function is the average amount of future reward expected when starting from each location.

To this end, in TD learning the time within a trial is explicitly represented (t below), and learning occurs at every timepoint within a trial, according to

$$V_{new}(S_i, t) = V_{old}(S_i, t) + \eta \Bigg[r(t) + \gamma \sum_{S_k @ t+1} V_{old} (S_k, t+1) - \sum_{S_k @ t+1} V_{old}(S_j, t) \Bigg] \qquad (22.2)$$

In this specific variant of TD learning, stimuli create long-lasting memory traces (representations), and a separate value $V(S, t)$ is learned for every timepoint of this trace (for instance, a stimulus can predict a reward 5 seconds after its presentation, but not 10 seconds later). As in the Rescorla-Wagner rule, η is a learning rate, and learning is driven by discrepancies between available and expected outcomes; however, the crux of the difference between the rules is in how predictions, or expectations, are construed. In TD learning, the associative strength of the stimuli (and traces) at time t is taken to predict not only the immediately forthcoming reward $r(t)$, but also the future predictions due to those stimuli that will still be available in the next time-step $\Sigma_{Ss@t+1} V(S, t+1)$, with $\gamma \leq 1$ discounting these future delayed predictions.

It turns out that the TD learning rule can be derived as a normative prediction learning rule. The formal justification for TD learning as a method for optimal reinforcement learning comes from its relation to dynamic programming methods (Sutton, 1988; Watkins, 1989; Barto *et al.*, 1990). Dynamic programming is a collection of computational methods for solving stochastic sequential decision problems (Bellman, 1957). Departing for the moment from animal conditioning and human decision making, consider a dynamic process (called a *Markov chain*) in which different states $S \in S$ follow one another according to some predefined probability distribution $P(S_{t+1} | S_t)$, and rewards are observed at each state with probability $P_r(S)$. A useful quantity to predict in such a situation is the expected sum of all future rewards, given the current state S_t, which we will call the *value* of state S

$$\begin{aligned} V(S_t) &= E[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \cdots | S_t] \\ &= E\left[\sum_{i=t}^{\infty} \gamma^{i-t} r_i \middle| S_t \right] \end{aligned} \qquad (22.3)$$

where $\gamma \leq 1$ is a factor discounting the effect of rewards distant in time on the value of the current state (this is necessary to ensure that the sum of future rewards is bounded). The expectation in equation (22.3) is with respect to two sources of stochasticity: (1) the probability of transitioning from one state to the next, and (2) the probability of reward in each state. Note that from this definition of state value it follows directly that

$$V(S_t) = P_r(S_t) + \gamma \sum_{S_{t+1}} P(S_{t+1} | S_t) V(S_{t+1}) \qquad (22.4)$$

This recursive relationship or *consistency* between state values lies at the heart of both dynamic programming and TD learning, and can be viewed as a kind of regularization or smoothing through time inasmuch as time indexes transitions from one state to another. The key to learning these values is that the recursive relationship holds *only* if the values are correct (i.e., they correctly predict the expected discounted sum of future values). If the values are incorrect, there will be a discrepancy between the two sides of the equation

$$\delta(t) = P_r(S_t) + \gamma \sum_{S_{t+1}} P(S_{t+1} | S_t) V(S_{t+1}) - V(S_t). \quad (22.5)$$

This *temporal-difference prediction error* is a natural "error signal" for improving estimates of the function $V(S_t)$ such that the discrepancy will be reduced

$$V(S_t)_{new} = V(S_t)_{old} + \eta \cdot \delta(t) \quad (22.6)$$

Returning to prediction learning in real-world scenarios, we note that this dynamic programming updating scheme has one problem: it requires knowledge of the dynamics of the environment, that is, $P_r(S)$ and $P(S_{t+1} | S_t)$ (the "world model") must be known in order to compute the prediction error δ(t) in equation (22.5). This is clearly an unreasonable assumption when considering an animal in a Pavlovian conditioning task, or a human predicting the trends of a stockmarket. Werbos (1977), in his "heuristic dynamic programming methods," and, later, Barto *et al.* (1989) and Bertsekas and Tsitsiklis (1996), suggested that in such a "model-free" case, the environment itself can supply this information stochastically and incrementally. An animal can *sample* the reward probabilities in each state, and the probabilities of transitions from one state to another, as it experiences the task. Updating according to these samples will eventually lead to the correct predictive values. Thus the stochastic prediction error

$$\delta(t) = r_t + \gamma V(S_{t+1}) - V(S_t) \quad (22.7)$$

(where r_t is the reward observed at time t, when in state S_t, and S_{t+1} is the next observed state of the environment) can be used as a Monte Carlo approximation to dynamic programming, in order to learn optimal predictive state values. The resulting learning rule

$$V(S_t)_{new} = V(S_t)_{old} + \eta(r_t + \gamma V(S_{t+1}) - V(S_t)) \quad (22.8)$$

is exactly the same rule as that proposed by Sutton and Barto (1990) in equation (22.2), if we add the Rescorla-Wagner-like assumption that the predictions of the different stimuli comprising the state of the environment are additive (which is not the only way to combine predictions, and is certainly not always the sensible option; see Dayan *et al.*, 2000). This shows that, using TD learning, animals can learn the optimal (true) predictive values of different events in the environment, even when this environment is stochastic and its dynamics unknown.

Optimal Action Selection

The discussion above holds whenever the probabilities of transitioning between different states or situations in the environment are stationary in time, as in Pavlovian conditioning (in which the animal cannot influence the events by means of its actions) or in situations in which the animal has a fixed behavioral policy (Sutton, 1988). But what about improving action selection in order to obtain more rewards – that is, what about instrumental conditioning? Since the environment rewards us for our actions, not for our predictions (be they correct as they will), the ultimate goal of prediction learning is to aid in action selection.

The problem of optimal action selection is especially difficult in those (very common) cases in which actions can affect long-term outcomes, or in which an outcome depends on a series of actions. For example, when winning or losing a game of chess, it is not at all simple to infer which were the actions responsible for this outcome, in order to improve the playing policy. This is true in the animal domain as well: when reaching a dead-end in a maze, how will a rat know which of its previous actions was the erroneous one? And, conversely, when it finds the cheese in the maze, how will it know which actions should be credited with the success? This is the (in)famous *credit assignment problem* (Sutton, 1978; Barto *et al.*, 1983; Sutton and Barto, 1998). RL methods solve the credit assignment problem by basing action selection not only on immediate outcomes but also on value predictions, such as those we discussed above, which embody long-term predictions of future outcomes.

First, note that, given predictive state values such as those learned by TD learning, a person could select the best long-term action at each state if only he knew what state that action would lead to (e.g., McClure *et al.*, 2003a). That is, given the transitions between states, the best action to choose is the one that leads to the state with the highest value. In fact, TD learning was first used in this way to select actions in Samuel's (1959) checker player. But what if this information is not available? For example, imagine deciding whether to buy or to sell a stock on the stock market – clearly, if you knew whether its price would increase or decrease as a result of your (and the rest of the market's) actions, this would be a trivial decision. But what can a human or a rat do in the completely model-free case – i.e., when it is not explicitly known how different actions will influence the state of the environment?

Actor/Critic methods

In one of the first RL papers (which was inspired by neural network models of learning), Barto *et al.* (1983) showed that the credit assignment problem can be solved by a learning system comprised of two neuron-like elements. One unit, the "adaptive critic element (ACE)," constructs an evaluation of different states of the environment, using a temporal-difference like learning rule from which the TD rule above was later developed. This evaluation is used to augment the external reinforcement signal and train a second unit, the "associative search element (ASE)," to select the correct action at each state through a trial-and-error process. These two elements were the precursors of the modern-day Actor/Critic framework for model-free action selection.

The insight in the ASE-ACE model, first due to Sutton (1978), is that even when the external reinforcement for a task is delayed (as in the game of chess), a temporal-difference prediction error can convey, at every time-step, a "reinforcement" signal to the action just chosen. In the absence of external reinforcement ($r_t = 0$), the prediction error $\delta(t)$ in equation (21.7) is equal to $\gamma V(S_{t+1}) - V(S_t)$; that is, it compares the values of two consecutive states. If the action has led to a state with a higher value than the previous state, this prediction error will be positive; if the situation has worsened due to the action taken, it will be negative. In the former case, the tendency to perform this action at this state should be increased (as the action has led to higher expected future rewards), while in the latter it should be decreased. Thus the agent can learn an explicit *policy* – a probability distribution over all available actions at each state $\pi(S, a) = p(a \mid S)$, and an adequate learning rule for this policy is

$$\pi(S, a)_{new} = \pi(S, a)_{old} + \eta_\pi \delta(t) \qquad (22.9)$$

where η_π is the policy learning rate and $\delta(t)$ is the prediction error from equation (22.7).

Thus, in Actor/Critic models, a critic module estimates state values $V(S)$ based TD learning from experience with the environment, and the same temporal-difference prediction error that is used to train the critic's values is also conveyed to the actor module, which maintains and learns a policy π (Figure 22.1). This method is closely related to policy improvement methods in dynamic programming (Sutton, 1988), and Williams (1992) has shown that in some cases the Actor/Critic can be construed as a gradient-climbing algorithm for learning a parameterized policy, which converges to a local minimum (see also Dayan and Abbott, 2001). However, in the general case Actor/Critic methods are not guaranteed to converge (*cf.* Baird, 1995; Konda and Tsitsiklis, 2003). Nevertheless, some of the strongest links between reinforcement learning methods and neurobiological data regarding animal and human decision making have been through the Actor/Critic framework. Specifically, Actor/Critic methods have been convincingly linked to action selection and prediction learning in the basal ganglia (e.g., Barto, 1995; Houk *et al.*, 1995; Joel *et al.*, 2002), as will be detailed in the next section[1].

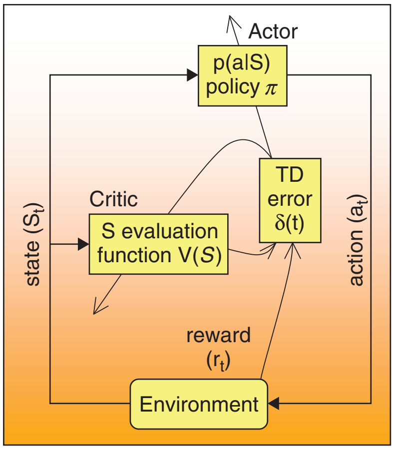


FIGURE 22.1 Actor/Critic architecture. The environment provides a state S_t and a reinforcement signal r_t to the critic, who uses these to compute a temporal difference prediction error (equation 21.7). This prediction error is used to train the state value predictions $V(S)$ in the critic as well as the policy $\pi(S, a)$ in the actor. Note that the actor does not receive information regarding the actual outcomes of its actions. Rather, the prediction-error signal is a proxy to these, telling the actor whether the outcomes are better or worse than expected. Figure adapted from Sutton and Barto, 1998.

State–action values

An alternative to Actor/Critic methods for model-free RL is to explicitly learn the predictive value (in terms of future expected rewards) of taking a specific action at a certain state. Building on dynamic programming methods of "policy iteration" (Howard, 1960), Watkins (1989) suggested *Q-learning*, a modified temporal-difference method in which the agent learns the value $Q(S, a)$ of each state-action pair (S, a) rather than

---

[1] More recent work (Morris *et al.*, 2006; Roesch *et al.*, 2007) suggests that the learned values may be more sophisticated and actually represent a value function defined over state–action pairs (rather than simply over states independent of actions), for which the learning algorithm is slightly different – see *Q*-learning or SARSA, see next section.

the value $V(S)$ of each state S. The learning rule itself is quite similar to the state value learning rule above

$$Q(S_t, a_t)_{new} = Q(S_t, a_t)_{old} + \eta\delta(t) \tag{22.10}$$

however, the temporal-difference prediction error term which drives Q-learning is slightly different

$$\delta(t) = r_t + \max_a \gamma Q(S_{t+1}, a) - Q(S_t, a_t) \tag{22.11}$$

where the $\max_a$ operator means that the temporal difference is computed with respect to what is believed to be the best available action at the subsequent state S_{t+1}. This method is considered "off-policy," as it takes into account the best future action, even if this will not be the actual action taken at S_{t+1}. In an alternative "on-policy" variant called SARSA (state–action–reward–state–action), the prediction error takes into account the actual chosen action, which leads to:

$$\delta(t) = r_t + \gamma Q(S_{t+1}, a_{t+1}) - Q(S_t, a_t) \tag{22.12}$$

In both cases, given such Q-values, action selection is easy, as the best action at each state S is that which has the highest $Q(S, a)$ value. Furthermore, dynamic programming results regarding the soundness and convergence of "policy iteration" methods (in which a policy is iteratively improved through bootstrapping of the values derived given each policy; Howard, 1960; Bertsekas and Tsitsiklis, 1996) ensure that if the proper conditions on the learning rate are met, these methods will indeed converge to the true optimal (in case of Q-learning) or policy-dependent (in the case of SARSA) state–action values.

APPLICATION OF REINFORCEMENT LEARNING MODELS TO NEURAL DATA

In recent years, RL models like those highlighted above have been applied to a wide range of neurobiological and behavioral data. In particular, the computational functions of neuromodulators such as dopamine, acetylcholine, and serotonin have been addressed using a growing family of RL models. Among these neuromodulatory systems, the dopamine system has long attracted the most attention, perhaps due to its well-known connection with disease processes like drug addiction, Parkinson's disease, and schizophrenia, as well as its role in reward learning and working memory. It is in elucidating the role of dopamine signals in the brain that computational models of learning in general, and TD learning in particular, have had their most profound and sustained impact on neuroscience.

Continuing in the spirit of a historical narrative, let us turn back two decades to the 1980s and early 1990s of the previous century, when it became clear that antipsychotic medication (i.e., dopamine receptor blockers), while mitigating many of the dramatically troubling symptoms of schizophrenia (auditory hallucinations, paranoia, etc.), also caused hedonic blunting. That is, patients receiving this type of medication appeared to not derive pleasure from stimuli and behavioral acts formerly known to cause pleasure. Dopamine receptor blockers were also shown to have detrimental effects on reward learning in laboratory animals. In light of these observations, Roy Wise proposed the "anhedonia hypothesis" of dopamine function (Wise *et al.*, 1978a, 1978b). According to this proposal, the role of dopamine is to mediate the rewarding or primary motivational characteristics of natural stimuli such as food, water, and sex, as well as those of drugs of abuse (Wise, 1982; for a recent review see Wise, 2004). Specifically, this suggested that dopamine equals reward; that is, that there is an equivalence between the level of dopamine in the brain and "reward value." Wise's hypothesis initiated a surge of studies into the effects of neuroleptics on reward-mediated learning, and the results indicated that blocking dopamine is like removing the reward contingent on an animal's actions (i.e., it causes extinction of responding, as if the reward is absent; see, for example, Franklin and McCoy, 1979; Willner *et al.*, 1987). The puzzle of the role of dopamine in the brain seemed close to being solved.

At that time, the lab of Wolfram Schultz, who pioneered single-unit recordings from the midbrain of awake and behaving monkeys, began recording the activity of dopaminergic neurons while monkeys underwent simple instrumental or Pavlovian conditioning (Romo and Schultz, 1990; Ljungberg *et al.*, 1992; Schultz *et al.*, 1993). As expected, these cells showed phasic bursts of activity when the monkey was given a rewarding sip of juice or a morsel of apple. Surprisingly, however, if food delivery was consistently preceded by a tone or a light, after a number of trials the dopaminergic response to reward disappeared. Contrary to the anhedonia hypothesis, the lack of measurable dopaminergic response to reward delivery did not accompany extinction, but rather acquisition – the monkeys began showing conditioned responses of anticipatory licking and arm movements to the reward-predictive stimulus. Indeed, not only the monkeys' responses to the tone, but also their dopaminergic neurons began responding to the tone, showing distinct phasic bursts of activity whenever the tone came on.

This pattern of results was also true for the difference between self-initiated reaching for reward (in which case dopamine neurons responded phasically to touching the reward) versus cue-initiated movements (in which case the neurons responded to the cue and not the reward). Rather than mediating the effects of affectively rewarding stimuli, it seemed that (quoting the conclusion sentences from papers of the time) "dopamine neurons exert a predominantly enabling effect on neurons more directly involved in the internal generation of movement" (Romo and Schultz, 1990: 592); "during acquisition of a simple behavioral task, dopamine neurons respond to unconditioned and conditioned salient stimuli that attract the attention of the animal, induce behavioral activation, and are associated with reward" (Ljungberg *et al.*, 1992: 145); and "dopamine neurons respond phasically to alerting external stimuli with behavioral significance whose detection is crucial for learning and performing delayed response tasks" (Schultz *et al.*, 1993: 900); all this "while not conveying specific information about the physical characteristics of stimuli nor the emergent behavioral reaction" (Romo and Schultz, 1990: 592).

A resolution of this conundrum was suggested in the mid 1990s, when Read Montague, Peter Dayan, Terry Sejnowski, and colleagues noticed that this pattern of dopaminergic responding throughout the course of learning conforms exactly to the characteristics of a reward prediction error (Montague *et al.*, 1993, 1994, 1995, 1996). Indeed, the hallmark of temporal-difference prediction errors is that they occur only when events are not predicted. For instance, in a simulated Pavlovian conditioning scenario in which a tone CS is followed two seconds later by a food US, prediction errors arise as a result of the unexpected US early in training when the relationship between the CS and the US is not yet known (Figure 22.2a), but not later in training when the CS comes to predict the US (Figure 21.2b). Providing that the CSs occur randomly and thus can not be predicted, at late stages of training they themselves generate a prediction error (similar to the one that had previously accompanied the US delivery) which can support second-order conditioning. Moreover, in trials in which the US is not delivered, a negative prediction error occurs at the precise time of the expected US delivery (Figure 21.2c; such precise timing also necessitates a stimulus representation that can track elapsed time, as detailed in the figure caption).

The close correspondence between the findings of Schultz and colleagues regarding phasic dopaminergic firing patterns and these characteristics of a temporal-difference prediction error (Figure 22.2d–f) led Montague *et al.* (1996) to suggest the *reward-prediction error theory of dopamine*[2] (see also Schultz et al., 1997). Within this theoretical framework, it was immediately clear why dopaminergic neurons fire to unexpected rewards but not to those that are predicted by previous stimuli, and why dopamine is necessary for reward-mediated learning in the basal ganglia. Indeed, the shift in dopaminergic activity from the time of reward to the time of the predictor (Takikawa *et al.*, 2004) resembles the shift of behavioral responses from the time of the US to that of the CS in Pavlovian conditioning experiments (Schultz *et al.*, 1997; Hollerman and Schultz, 1998). Moreover, the model explained why, after training, dopamine neurons did not fire above baseline in the time period between a predictive cue and the reward delivery – in the absence of new information, there are no prediction errors at these intermediate times. From the point of view of downstream neural structures, continuous baseline firing of dopaminergic neurons can be interpreted as "things are just as expected."

The basic characteristics of phasic dopaminergic responding have since been replicated in many variants (Hollerman and Schultz, 1998; Schultz, 1998; Tobler *et al.*, 2003; Takikawa *et al.*, 2004; Bayer and Glimcher, 2005). In fact, recent work aimed at putting the prediction-error hypothesis to quantitative test has demonstrated that the correspondence between phasic dopaminergic firing and temporal-difference prediction errors goes far beyond the three basic characteristics depicted in Figure 21.2. For instance, using a general regression model that does not assume temporal-difference learning *a priori*, Bayer and colleagues (Bayer and Glimcher, 2005; Bayer *et al.*, 2007) have rigorously shown that the contribution of previously experienced rewards to the dopaminergic response to the reward in the current trial is exactly according to an exponentially weighted average of past experience (as is implicit in the temporal-difference learning rule; see Figure 22.3). Moreover, conditioned stimuli predicting probabilistic rewards or rewards of different magnitudes have been shown to elicit a phasic dopaminergic response that is indeed proportional to the magnitude and/or probability of the expected reward (Fiorillo *et al.*, 2003; Morris *et al.*, 2004; Tobler *et al.*, 2005; Figure 22.4a, b), and firing patterns in tasks involving probabilistic rewards are in accord with a constantly back-propagating error signal (Niv *et al.*, 2005b; Figure 22.4b, c). Even in sophisticated

---

[2] Interestingly, dopaminergic neurons do not seem to be involved in the signaling or prediction of aversive stimuli (Mirenowicz and Schultz, 1996; Tobler *et al.*, 2003; Ungless *et al.*, 2004), in which the neuromodulator serotonin has been implicated instead (Daw *et al.*, 2002).

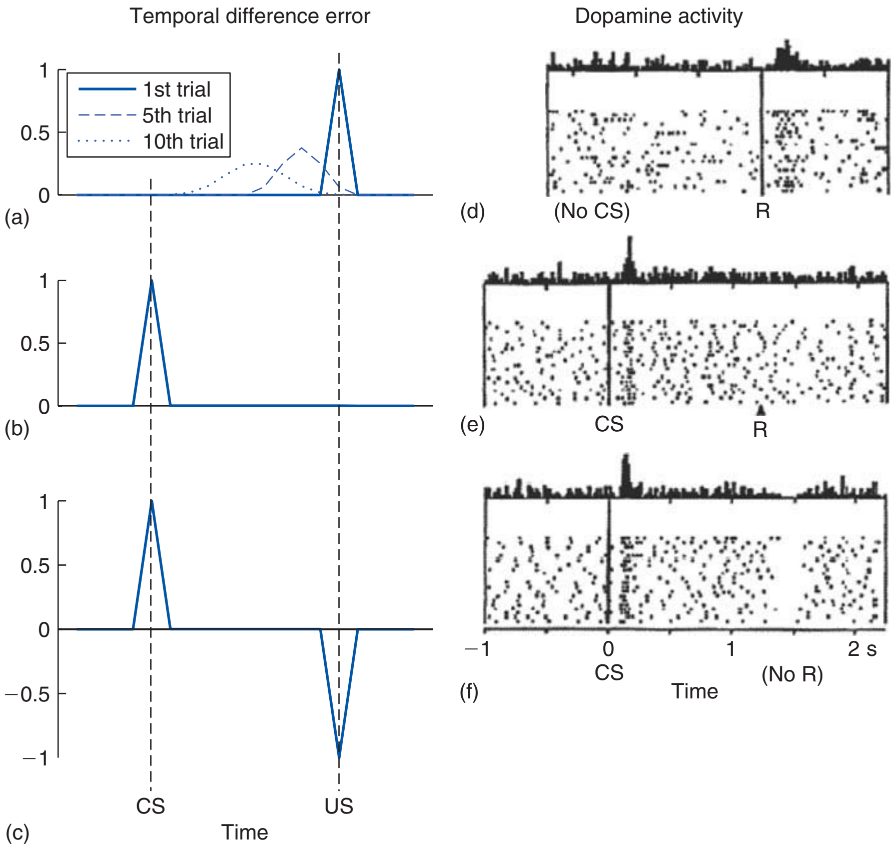


FIGURE 22.2 (a–c) Temporal-difference prediction errors in a simple Pavlovian conditioning task. A tone CS is presented at random times, followed two seconds later with a food US. At the beginning of training (a), the affectively significant US is not predicted, resulting in prediction errors. As learning progresses (trials 5 and 10 are plotted as examples), the prediction error propagates back (Niv *et al.*, 2005a) as values of preceding timesteps are updated (equation 21.8). When the predictive relationships are completely learned (b), the now-predicted US no longer generates a prediction error, rather, the unpredicted occurrence of the CS is accompanied by a prediction error. If the US is unexpectedly omitted (c), a negative prediction error is seen at the time in which the US was expected, signaling that expectations were higher than reality. In these simulations, the CS was represented over time with the commonly used serial compound state representation (Kehoe, 1977, Sutton and Barto, 1990), and there was no discounting ($\gamma = 1$). (d–f) Firing patterns of dopaminergic neurons in the ventral tegmental areas of monkeys performing an analogous instrumental conditioning task. Each raster plot shows action potentials (dots) with each row representing a trial, aligned to the time of the cue (or the reward). Bar histograms show the summed activity over the trials plotted below. When a reward is given unexpectedly, dopaminergic neurons respond with a phasic burst of firing (d). However, after conditioning with a predictive visual cue (which, in this task, predicted a food reward if the animal quickly performed the correct reaching response), the predicted reward no longer elicits a burst of activity, and the phasic burst now accompanies the presentation of the predictive cue (e). In "catch" trials, in which the food reward was unexpectedly omitted, dopaminergic neurons showed a precisely-timed pause in firing, below their standard background firing rate (f). Subplots 22.2 (d–f) adapted from Schultz *et al.*, (1997). Note that the discrepancies between the simulation and the dopamine neuron firing patterns in terms of the magnitude and spread of the prediction errors at the time of the reward likely result from the temporal noise in reward delivery in the instrumental task, and the asymmetric representation of negative and positive prediction errors around the baseline firing rate of these neurons (Niv *et al.*, 2005a).

conditioning tasks such as blocking and appetitive conditioned inhibition, Waelti *et al.* (2001) and Tobler *et al.* (2003, 2005) have shown that the dopaminergic response is in line with the predictions of temporal-difference learning. And in the timing domain, recent results show that the dopaminergic activity to a cue predicting a delayed reward is attenuated in proportion to the delay (Figure 22.5), as is expected from a signal predicting the expected sum of *discounted* future rewards (Roesch *et al.*, 2007). Finally, direct measurements of extracellular dopamine in the nucleus accumbens (a brain area also referred to as ventral striatum; a major target for dopaminergic projections from the ventral tegmental area) using fast-scan cyclic voltammetry (which has subsecond temporal resolution) have confirmed that the phasic changes in levels of dopamine in target structures indeed conform quantitatively to a prediction error signal (Paul Phillips, personal communication; see also Day *et al.*, 2007; Chapter 24 of this volume), despite the non-linearities between dopamine neuron firing and actual synaptic discharge of the transmitter (Montague *et al.*, 2004).

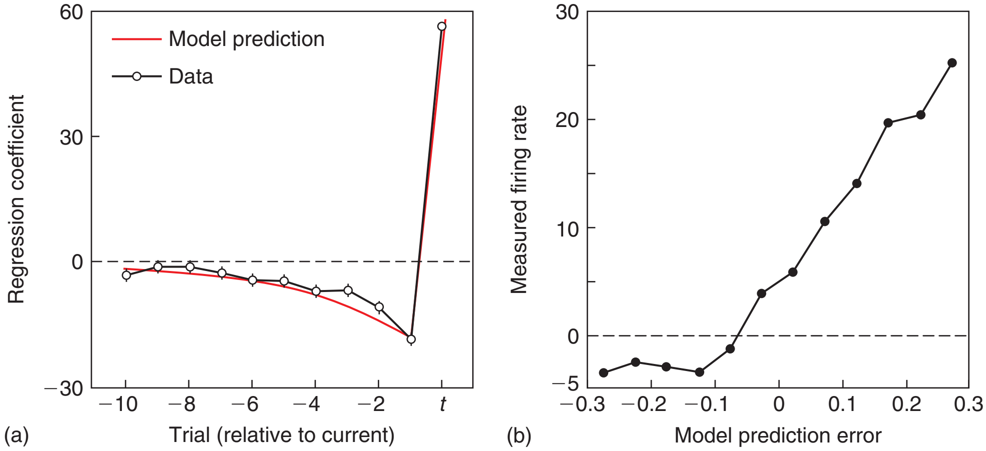


FIGURE 22.3 Dopaminergic responses depend on past rewards, as is predicted by temporal-difference learning. Here, single-unit recordings of dopaminergic neurons were conducted while monkeys performed a rewarded saccade task (Bayer and Glimcher, 2005; Bayer *et al.*, 2007). In this task, a visual cue signaled the start of a trial, after which the monkey could choose when to perform a saccade toward a target. The magnitude of the reward (volume of fruit juice) depended on the time of the saccade, such that a longer wait was rewarded with more juice, up to an unsignaled deadline, after which saccades earned no juice. This allowed for a range of prediction errors to be sampled. The monkey's task was to learn the optimal time-to-look in order to receive the maximum reward. (a) Prediction errors at the time of the reward were regressed against the rewards in the past 10 trials, which should, according to TD learning, determine the amount of reward predicted (and thus the magnitude of the prediction error). In solid red is the dependency of the prediction error on past rewards as derived from the theoretical model (with a learning rate of $\eta = 0.7$; see Bayer and Glimcher, 2005 for details), in black circles are the regression coefficients obtained from the data. The data are in close fit with the model prediction. (b) the measured prediction error is linear with the model-derived prediction error, at least in the domain of positive prediction errors (see Bayer *et al.*, 2007, for data indicating that negative prediction errors are encoded by the length of the pause in firing, rather than the magnitude of the pause below the baseline firing rate). Figure adapted from Bayer and Glimcher, 2005.

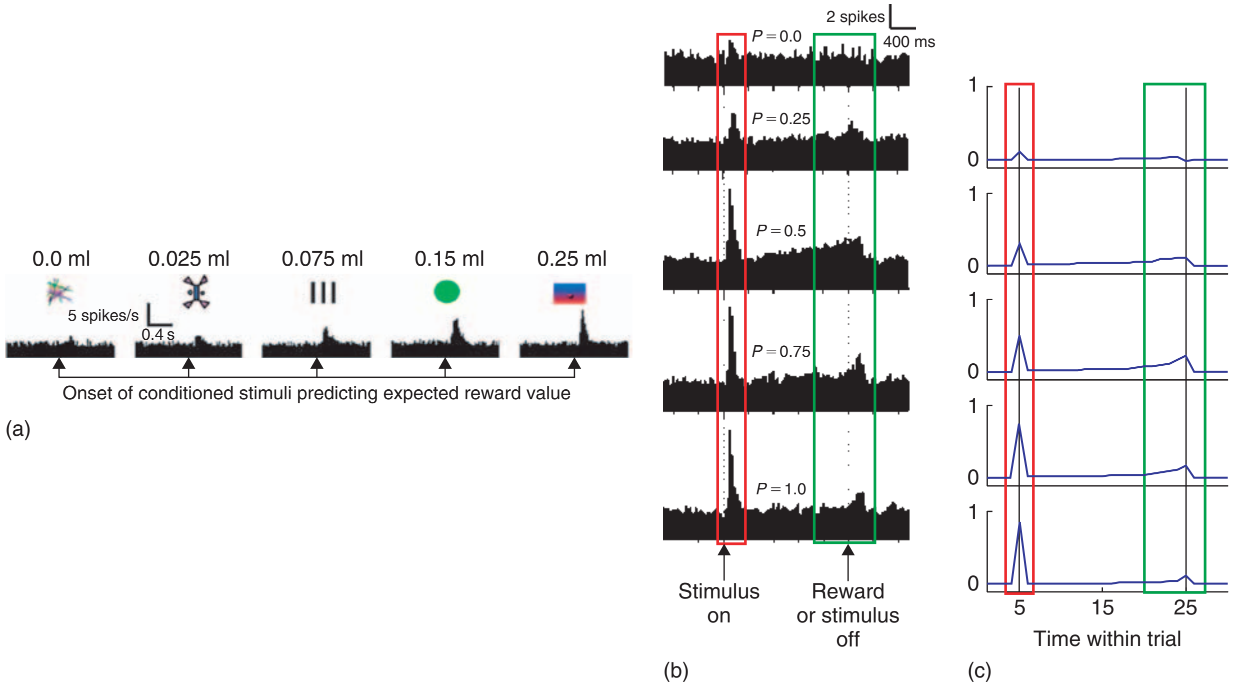


FIGURE 22.4 Dopaminergic responses comply with the predictions of temporal-difference learning. (a) Phasic prediction errors at the time of a cue predicting reward are proportional to the magnitude of the predicted reward (adapted from Tobler *et al.*, 2005). (b, c) When different cues predict the same reward but with different probabilities, the prediction error at the time of the cue is proportional to the predicted probability of reward (red rectangles; compare panel (b) (data) to panel (c) (TD simulation)). However, due to the low baseline firing rate of midbrain dopaminergic neurons, negative prediction errors cannot be encoded with as deep a "dip" in firing rate as is the height of the "peak" by which positive prediction errors are encoded. As a result, when rewards are probabilistic, averaging over rewarded and unrewarded trials will create an apparent ramp leading up to the time of the reward (green rectangles; compare panel (b) (data) to panel (c) (TD simulation)). Panel (b) adapted from Fiorillo *et al.*, 2003; Panel (c) adapted from Niv *et al.*, 2005a.

Note that the prediction-error theory of dopamine is a *computationally precise* theory of the *generation* of phasic dopaminergic firing patterns. It suggests that dopaminergic neurons combine their diverse afferents (which include inputs from the medial prefrontal cortex, the nucleus accumbens shell, the ventral pallidum, the central nucleus of the amygdala, the lateral hypothalamus, the habenula, the cholinergic pedunculopontine nucleus, the serotoninergic raphe, and the noradrenergic locus coeruleus; Christoph *et al*., 1986; Floresco *et al*., 2003; Geisler and Zahm, 2005; Matsumoto and Hikosaka, 2007; Kobayashi and Okada, 2007) to compute a temporal-difference reward-prediction error. Moreover, it suggests that dopamine provides target areas with a neural signal that is theoretically appropriate for controlling learning of both predictions and reward-optimizing actions. Following the analogy between the dopamine signal and the temporal-difference prediction-error signal in Actor/Critic models (Joel *et al*., 2002), it has been suggested that the signal reported by dopaminergic neurons in the ventral tegmental area to ventral striatal and frontal target areas, is used to train predictions (as in the critic; Barto, 1995; Waelti *et al*., 2001), while a similar signal reported by dopaminergic neurons in the substantia nigra pars compacta to dorsal striatal target areas is used to learn an action-selection policy (as in the actor; Miller and Wickens, 1991; Wickens and Kötter, 1995; Houk *et al*., 1995; Joel and Weiner, 1999).

Recently, the role of phasic dopamine in action selection was assessed by combining non-trivial decision-making tasks in rodents or monkeys with single-cell recordings of dopaminergic activity. This is especially interesting, as temporal-difference methods do not fully prescribe the form of the reward-prediction error in tasks involving action selection. As mentioned above, different computational algorithms, namely Actor/Critic, *Q*-learning and SARSA, make different predictions regarding the nature of the cue-related prediction error (Niv *et al*., 2006b), making electrophysiological evidence critical in constraining the algorithm actually used by the brain. In a recent study, Morris *et al*. (2006) trained monkeys with cues predicting reward with different probabilities. Interestingly, in "catch" trials, in which the monkeys were given a choice between two cues, the cue-elicited prediction errors matched best the errors corresponding to the cue that would subsequently be chosen. This is contrary to the straightforward predictions of an Actor/Critic mechanism, and more in line with SARSA learning. However, data from rats performing a more dynamic odor-discrimination task (Roesch *et al*., 2007) and from monkeys engaged in a difficult random-dot motion discrimination task (Nomoto *et al*., 2007) suggest that predictions (and thus prediction errors) can be sensitive to the information available at every timepoint, representing stimuli before a choice is made, and representing the chosen cue only later. These different results can be incorporated into one learning scheme using appropriate representation of the states in the task, an issue that we shall return to in the last section of this chapter.

To end this section, we should mention that there are alternative psychological theories regarding the role of dopamine in conditioned behavior (for a recent debate-style review, see Berridge, 2007). These include Redgrave and colleagues' "incentive salience" (see, for example, Redgrave *et al*., 1999; Horvitz, 2000; Redgrave and Gurney, 2006), Berridge and

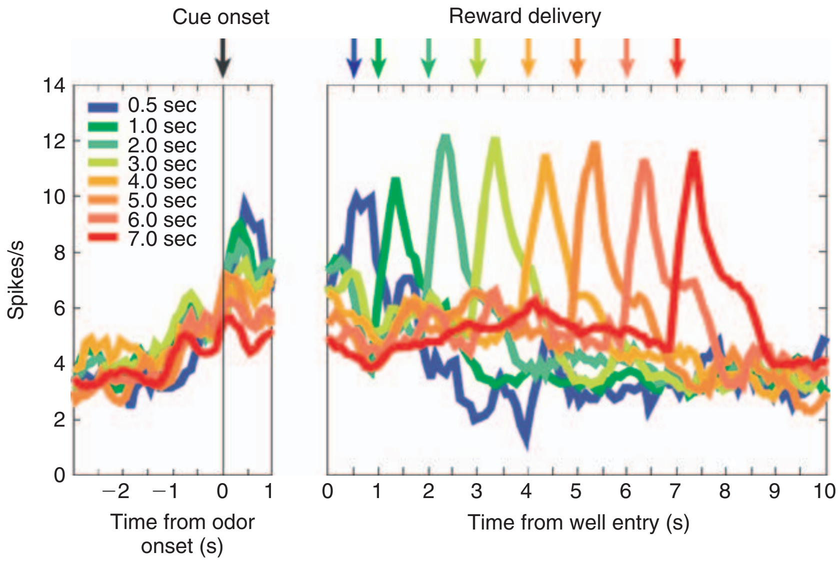


FIGURE 22.5 Average firing rate of 19 dopaminergic neurons, recorded in rats performing an odor-discrimination task in which one of the odors predicted a delayed reward that will be delivered in a food-well. Color indicates the length of the delay preceding reward delivery from 0.5 to 7 seconds. Activity is aligned on odor onset (left) and food-well entry (right). Rewards were given in the well providing that the rat entered the correct well (as indicated by the odor cue), and remained there until the time of the reward. Adapted from Roesch et al. (2007).

Robinson's "wanting" versus "liking" (e.g., Berridge and Robinson, 1998; Berridge, 2007), and ideas about dopamine signaling uncertainty (Fiorillo *et al.*, 2003). A discussion of the merits and pitfalls of the different theories is beyond the scope of this chapter, and in many cases would involve the less-than-satisfactory comparison of qualitative suggestions to quantitative predictions of an RL model. Nevertheless, in as far as these theories are indeed fundamentally different from the prediction-error theory (which is not always clear), it is our opinion that, to date, no alternative has mustered as convincing and multidirectional experimental support as the prediction-error theory of dopamine.

EVIDENCE FROM IMAGING OF HUMAN DECISION MAKING

Although animal conditioning can display complex phenomena that are still well beyond our current understanding of the brain, ultimately we are interested in understanding human decision making, the computations that underlie it, and the relationship of these computations to neural mechanisms. While the characteristics of human conditioning are similar to those of animal conditioning, the possibility of instructing subjects verbally allows for much more elaborate paradigms in human experiments. Of course, there are severe limitations on our ability to measure neural processes in humans. One technique that has recently been brought to the forefront is functional magnetic resonance imaging (fMRI), in which metabolic correlates of neural activity can be measured non-invasively, albeit at low temporal and spatial resolution (seconds and millimeters, respectively), and subject to noisy measurements (typically necessitating averaging over trials or subjects).

One great advantage of fMRI is that it allows imaging of activity throughout the entire brain, rather than in only a small population of neurons. Using fMRI as the neural probe of choice places a premium on using precise computational models of decision making. A model-driven analysis also affords a special insight into brain function: the models identify *hidden variables* that control behavior, such as state values or prediction errors, for which we can search in the brain. This is different from the more straightforward identification of the neural correlate of perception or motor behavior, and even goes beyond a search for the correlates of abstract entities (such as "reward" or "preference;" e.g., O'Doherty *et al.*, 2002) because the computational model can quantitatively specify the dynamics of a hidden variable within a non-stationary learning and decision-making task. Identifying a neural correlate of such a signal advances our understanding of the brain in a way that would not be possible without the model. Of course, this can also lend powerful support for the model that gave rise to the specific values of the hidden variable: models that suggest different relationships between external events (stimuli and rewards) and internal representations (values and preferences) can be compared by testing how well activity in different brain areas correlates to the specific predictions of each model. With these general issues in mind, we turn below to the specific use of RL models in identifying learning signals (e.g. reward-prediction errors) and value-dependent responses in the human brain.

The first fMRI studies to search for prediction errors in humans implicated the nucleus accumbens and the orbitofrontal cortex (Berns *et al.*, 2001; Knutson, *et al.*, 2001a; Pagnoni *et al.*, 2002), both major dopaminergic targets. O'Doherty *et al.* (2003) and McClure *et al.* (2003b) then used a hidden-variable analysis technique to identify the neural correlates of model-derived temporal-difference prediction errors. These studies again implicated the nucleus accumbens (the ventral striatum) and the putamen (part of the dorsal striatum). Later, functional imaging was used to distinguish between potential sites of Pavlovian versus instrumental learning: O'Doherty *et al.* (2004) showed that fMRI correlates of prediction-error signals can be dissociated in the dorsal and ventral striatum according to whether an action is required in order to obtain reward. For passive prediction-learning tasks the reward-prediction error was evident only in the ventral striatum, while in active tasks it was evident in both the ventral and the dorsal striatum (Figure 22.6). These findings and the model-based analysis that uncovered them suggest that stimulus–response learning typical of Actor/Critic circuits in humans may be associated with activation in the dorsal striatum.

Indeed, correlates of prediction errors in the dorsal and ventral striatum have now been seen in multiple studies (see, for example, Li *et al.*, 2006; Preuschoff *et al.*, 2006; Schönberg *et al.*, 2007). Note, however, that the fMRI results cannot isolate dopaminergic activity from other activity in the brain. Furthermore, the measured blood oxygen level dependent (BOLD) signal in a brain area has been suggested to be correlated with local field potentials[3] implying a correlation with the afferent inputs to and the local processing in a brain region (Logothetis, 2003), rather than the spiking activity of neurons within the region. Thus, prediction-error correlates in areas of the striatum and the prefrontal cortex are fully in line with the fact that

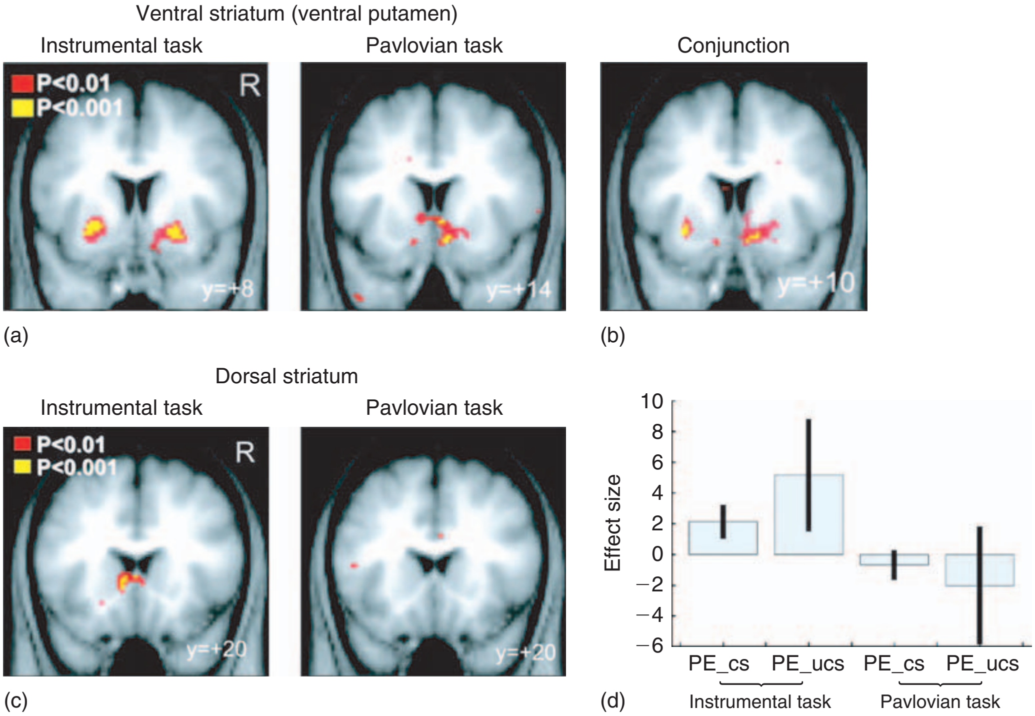


FIGURE 22.6 BOLD correlates of reward-prediction error in a Pavlovian and an instrumental task can be anatomically dissociated. (a) Correlates of a reward-prediction error signal are seen in the ventral striatum (specifically, in the ventral putamen) in both an instrumental task involving action selection in order to maximize rewards, and a yoked Pavlovian task in which the subject observes choices made by the computer. (b) The conjunction of the activations verifies a similar involvement of the ventral striatum in both task variants. (c) However, correlates of reward-prediction errors are seen in the dorsal striatum only in the instrumental task. (d) Coefficient estimates for the prediction error at the time of the stimulus (PE_cs) and the reward (PE_ucs) for each task, from the peak voxel for the contrast: instrumental prediction error > Pavlovian prediction error. Figure adapted from O'Doherty *et al.*, 2004.

these are the major targets for dopaminergic influence. Furthermore, dopaminergic manipulations (e.g., administration of dopamine enhancers (agonists) or dopamine receptor blockers (antagonists)) in such tasks have been shown to influence both the BOLD measurement of prediction-error activity and learning and action selection (Pessiglione *et al.*, 2006), and recent results show that better learners show a higher correlation of striatal BOLD with a reward-prediction error (Schönberg *et al.*, 2007). Of course, these areas are also targeted by other afferents, most notably the neuromodulator *serotonin*, which has been suggested as the counterpart to dopamine in the domain of punishment (Daw *et al.*, 2002), and might explain why BOLD correlates of *positive* prediction errors for pain and punishment have also been found in the striatum (Seymour *et al.*, 2004; Jensen *et al.*, 2007; Menon *et al.*, 2007). The relative contribution of many neural signals to the measured BOLD responses in these regions awaits more precise pharmacological manipulations and perhaps a serious technological advance.

Note also that, without temporal specificity and an analysis that specifically aims to tease apart different components of the reinforcement learning model, it is not easy to distinguish between prediction errors and state values at the time of a stimulus. This is because the prediction error at the time of an unpredicted stimulus is $\delta(t) = V(stimulus) - V(baseline)$, which, if we take $V\ (baseline)$ to be 0, is exactly equal to $V(stimulus)$. Indeed, many studies have implicated the striatum in representing the anticipated value of outcomes (e.g., Knutson *et al.*, 2001a; Delgado *et al.*, 2003; Knutson *et al.*, 2003), and it is not always clear whether the measured activation is distinct from that attributable to a prediction error. In any case, electrophysiological data show that the striatum is definitely a viable candidate for representing state values (e.g., Schultz *et al.*, 1992; Samejima *et al.*, 2005). Studies in which outcomes involved both gains and losses have further implicated the striatum in the anticipation of losses, not only gains, with a decrease in activity correlated with the

---

[3] Local field potentials are electrophysiological signals that are related to the sum of all dendritic synaptic activity within a volume of tissue, thus they are dominated by the arrival of action potential along axons that terminate in the area, rather than the firing of neurons in that area. The local field potential is believed to represent the synchronized input into the observed area, as opposed to the spike data, which represent the output from the area.

anticipated loss. Moreover, the degree of deactivation to losses compared to activation to gains ("neural loss aversion") in the nucleus accumbens and the prefrontal cortex was predictive of individual differences in behavioral loss aversion (Tom *et al.*, 2007). Finally, outcome values themselves (as well as subjective preferences) have been associated with activations in areas such as the ventromedial prefrontal cortex and the orbitofrontal cortex (e.g., Knutson *et al.*, 2001b; O'Doherty *et al.*, 2002; Knutson *et al.*, 2003; McClure *et al.*, 2004).

The promise of model-driven analysis of imaging data has yet to be fully realized, and the link between computational models of learning and the brain does not end with the identification of the reward-prediction error signal. Recent work has used such a hidden-variable analysis coupled with a reinforcement learning model to investigate the neural substrates of exploration (Daw *et al.*, 2006). In "market-like" tasks, model-based approaches have identified learning signals related to so-called fictive outcomes (what might have happened but didn't, also called countefactuals; Lohrenz *et al.*, 2007) and a hierarchical RL model has been used to demonstrate that the brain tracks the volatility (or rate of change) of the environment (Behrens *et al.*, 2007). Furthermore, contrasting model-free temporal-difference learning with model-based learning algorithms that exploit the higher order structure of the learning task, Hampton *et al.* (2006) have begun to reveal the neural mechanisms of more model-based forms of human learning. One approach that is becoming increasingly common is the use of functional imaging in combination with pharmacological challenges (e.g., Pessiglione *et al.*, 2006) or with radioligand-labeled positron emission tomography (e.g., Zald *et al.*, 2004) to test more directly the causal predictions and pharmacological hypotheses of reinforcement learning models in human subjects (Knutson and Gibbs, 2007), respectively.

BEYOND PREDICTION ERRORS AND PHASIC DOPAMINE

The theoretical importance of prediction errors and prediction learning is undeniable. However, other computationally important signals have been posited and associated with neural function. In this section, we briefly discuss other components of learning beyond those of prediction errors, and how these might be related to different aspects of neuromodulation in the brain. We begin by applying the reinforcement learning framework to decisions about how fast (or with what vigor) to behave. This highlights the role played by the passage of time (specifically, the opportunity cost of time in terms of devoting time to one action rather than another) in ongoing decision making. We focus on a recent model that shows that the net rate of rewards quantifies the opportunity cost of time, and discuss the proposal that this quantity is represented by tonic levels of dopamine in the striatum (Niv *et al.*, 2007; Niv, 2007a). Dopamine has been argued to convey (at least) two separate signals to efferent (downstream) structures (e.g., Grace, 1991; Floresco *et al.*, 2003): On the one hand, firing of dopaminergic neurons induces large but short-lived changes in the amount of dopamine in a synapse (a "phasic" signal, predominantly affecting D1-type low-affinity dopamine receptors). On the other hand, extrasynaptic levels of dopamine change on a much slower timescale (a "tonic" signal which affects high affinity D2-type dopamine receptors). The computational and theoretical differences between phasic and tonic aspects of prediction highlight the importance of carefully treating different timescales of dopamine signaling, and suggest that dopamine can simultaneously fulfill several roles in decision making, without contradiction.

In the second part of this section, we discuss normative adaptations of the rate of learning to the decision-making task and the statistics of the environment. Bayesian inference models such as the Kalman filter show how different forms of uncertainty about the environment should affect the rate of learning, and the degree of reliance on previous experience. These effects have recently been associated with both acetylcholine and norepinephrine and their effects on learning and inference (Yu and Dayan, 2002, 2005), considerably enhancing our understanding of the neural basis of learning.

Tonic Dopamine and the Choice of Response Vigor

It is somewhat of a curiosity that although the tradition in animal experimentation is to investigate the determinants of *rates* of responding (as in Skinner's investigations of key-pecking in pigeons or lever-pressing in rats, so called "free-operant" experiments because the animal is free to choose when to respond and no trial structure is imposed on behavior), reinforcement learning models of conditioning have concentrated exclusively on the choice of discrete actions at pre-specified timepoints (as in a discrete trial two-alternative choice experiment). However, real-life decisions most often take place in continuous time.

In fact, every choice of action, even that in a discrete trial setting, is accompanied by a choice of the *speed* or *vigor* with which that action will be performed. This decision gives rise to response rates in free operant behavior, to running times in mazes, and to reaction-time data in discrete settings. It also interacts with the influences of motivation on behavior – a hungry rat running down a maze in search of food will run faster than a sated rat.

That reinforcement learning has, until recently, wholly ignored this aspect of decision making may be due to the historical origins of reinforcement learning theory in computer science. In simulations, or in robot decision making, decisions can only occur in synchrony with a discrete CPU clock. However, theory does exist that deals with continuous time: this is *average reward* reinforcement learning in a *semi*-Markov decision process (Schwartz, 1993; Doya, 2000; Daw and Touretzky, 2002). Building on this theoretical framework, Niv *et al.* (2005a) recently proposed a reinforcement learning model of optimal rates of responding. In this model of instrumental conditioning, every choice of action is accompanied by a choice of a *latency* with which to perform that action, such that the net overall rate of rewards is maximized. The model successfully replicates and explains the fundamental characteristics of free operant response rates (Niv, 2007b), and explains how motivational states should affect decision making (Niv *et al.*, 2006a).

Importantly, the average reward framework highlights an important factor that determines optimal responding: the net rate of rewards, that acts as the opportunity cost of time. To illustrate this, imagine a rat pressing a lever in order to obtain food. Suppose that through previous behavior food had been accrued at a rate of four pellets per minute. When contemplating devoting five seconds to executing the next lever-press, the potential benefit of this action (i.e., the probability of its generating reward, and the magnitude of this reward) should thus be weighed against both the (motor and other) costs of performing the action at this speed, and the opportunity cost of time, i.e., the potential loss of (on average) ⅓ reward pellets due to devoting time to this action rather than continuing to behave according to the previous policy. Because of this cost/benefit tradeoff, the model predicts that when the net rate of rewards is higher all actions should optimally be performed faster, as a result of the elevated opportunity cost (Niv *et al.*, 2007).

How does this relate to decision making in the brain? Note that the prediction-error theory of dopamine concentrates on only one aspect of dopaminergic activity and influence: the effect of *phasic* dopaminergic signaling on learning and plasticity. However, dopamine neurons operate in both a phasic and a tonic mode (Grace, 1991; Weiner and Joel, 2002; Bergstrom and Garris, 2003; Floresco *et al.*, 2003; Goto and Grace, 2005), and affect not only synaptic plasticity, but also membrane potentials and neural excitability, which may be particularly sensitive to tonic levels of dopamine (Nicola *et al.*, 2000; Schultz, 2002). Furthermore, the effects of dopaminergic manipulations such as lesions, antagonism, or agonism, are first and foremost seen in the vigor of ongoing behavior, rather than in learning processes. For instance, a multitude of studies has shown that injections of 6-hydroxydopamine into the nucleus accumbens, which causes the death of dopaminergic neurons projecting to that area, profoundly reduce the rate of instrumental responding (for a review, see Salamone and Correa, 2002). As a result, dopamine in the striatum has been linked to invigorating Pavlovian and instrumental responding (Ikemoto and Panksepp, 1999; Salamone and Correa, 2002).

Combining these lines of evidence, Niv and colleagues have suggested that tonic levels of striatal dopamine represent the net rate of rewards[4]. This hypothesis, dovetailing neatly with both computational theories regarding phasic dopamine signals and appetitive prediction errors, and psychological theories about dopamine's role in energizing responses, provides the first normative explanation for the critical role that tonic levels of dopamine play in determining the vigor of responding. It also suggests a route by which dopamine could mediate the effects of motivation on response vigor.

Acetylcholine and Norepinephrine and the Optimal Rate of Learning

One issue that we have not yet discussed relates to the assumption, in both the Rescorla-Wagner model and the temporal-difference model, that the predictions of multiple stimuli are simply added up to form the total prediction of reward. At this point, we can treat this as but a simplification: even the Rescorla-Wagner model allowed for different learning rates for different stimuli (for instance, based on their salience, as in overshadowing), implying that the prediction error should not affect all stimuli equally. A natural extension of this idea is to allow for stimuli to contribute

---

[4] In fact, if the tonic level of dopamine reflects spillover from phasic prediction error signals averaged over a longer time-frame due to slow reuptake, it follows computationally that it would, by default, equal the net rate of obtained rewards.

differentially to the overall prediction itself. Here again, control theory can inform us regarding the optimal combination of predictions and allocation of learning to different predictors. Simply put, unreliable stimuli, those with which we have had less experience and thus know less about, should contribute less to the prediction. As in the more general statistical problem of combining multiple sources of evidence, this implies a competitive interaction between predictors, with the most reliable predictor weighted most heavily in an overall weighted average. Conversely, when a prediction error occurs, more learning should be devoted to the stimuli about which there is most uncertainty, that is, they should take responsibility for most of the prediction error (Dayan and Kakade, 2000; Dayan *et al*., 2000).

These ideas are formally couched in statistically optimal Bayesian inference in the Kalman filter model. The Kalman filter assumes an underlying generative model in which each stimulus gives rise to observations (rewards) distributed according to a Gaussian distribution. Furthermore, it assumes that the process by which the rewards are observed (or measured) is prone to Gaussian noise. Different from Rescorla-Wagner or temporal difference learning, the learning process in this case must infer the mean reward associated with each stimulus while taking into account these two sources of variability. Optimal Bayesian inference dictates tracking not only the mean predicted reward, but also the uncertainty in this mean as observations accumulate (thus the learning process includes two learning or update rules). However, the optimal learning rule for the mean is rather similar to the Rescorla-Wagner learning rule with an additional adjustment of the learning rate based on the different sources of variance and on the tracked uncertainty (Dayan *et al*., 2000).

Yu and Dayan (2005) have further analyzed the effects of uncertainty on learning and inference in a noisy and constantly changing environment. Their model accounts for two types of uncertainty: *expected* uncertainty that arises from known variability in predictors in the environment, and *unexpected* uncertainty due to sudden unforeseen changes in the environment. Building on physiological, pharmacological, and behavioral data, they hypothesize that the first source of uncertainty is signaled by acetylcholine and the second by norepinephrine. Inference in their model is considerably more complex than in the case of the Kalman filter, and approximations are needed to render it feasible for a neural mechanism. In terms of neuromodulation, while acetylcholine and norepinephrine are synergistic in the learning process, their relationship in inference is antagonistic – inconsistencies between prediction and observation must be attributed either to expected or to unexpected uncertainty. If the former is low, the inference will be of an unexpected change in the environment; conversely, when the environment is deemed very stable, prediction errors imply higher (expected) variability of the predictors.

WHAT'S MISSING? CHALLENGES AND FUTURE DIRECTIONS

RL models are now used routinely to design and interpret a wide range of reward learning and decision-making experiments; however, we view this success only as an important starting point. One of the reasons that RL models have been successful is that they have been made extremely simple and have been tested in very simple experimental settings. In this last section, we point out some of the experimental data that challenge RL models and separate what we consider to be real challenges from confusion arising from a lack of clarity about the nature of the claims made by the modeling efforts.

The first challenge emerges from a range of responses of dopamine neurons to stimuli not clearly related to reward prediction. For example, novel stimuli have been shown to cause phasic bursts in dopamine neurons (Schultz, 1998) including even nociceptive stimuli (Coizet *et al*., 2006). By virtue of being novel, such stimuli should not be predictive of any outcome, aversive or appetitive. However, learning and perception are not done on the background of a blank slate. It is reasonable to think that generalization to previously encountered stimuli would play a critical role in the initial appraisal of a novel stimulus. If the experimental (or the general ecological) scenario is such that animals have learned to expect that stimuli predict rewards (as is the case in many experimental situations), it is not surprising that new stimuli will be treated optimistically. Kakade and Dayan (2002) directly addressed this possibility, and furthermore suggested that the novelty responses can function as novelty bonuses that enhance exploration of novel stimuli. In this account, novelty itself acts as a reward and combines with current reward information r_t to form a reward signal sensitive to novelty, that is, $r_{new}(t) = r_t + novelty(S_t)$. Kakade and Dayan show how this simple maneuver accounts in detail for the reported novelty responses of dopamine neurons (for instance, for the observation that the novelty burst is frequently followed immediately by a dip of the firing rate below baseline) yet still explains how they also communicate a reward prediction error.

One would hope that this rather innocuous-looking change in the model did not change appreciably what it learns. In fact, as shown by Ng *et al.* (1999), it does not, and Kakade and Dayan demonstrate clearly how to apply these results to the anomalous dopaminergic data. One issue brought to the forefront by this work is whether the dopamine system responds to aversive stimuli, and whether there should be an opponent system interacting with the information encoded in dopaminergic activity. This possibility is currently under development (Daw and Touretzky, 2000; Daw *et al.*, 2002; Doya, 2002).

A second challenge for RL models in explaining dopaminergic function has arisen primarily due to the complaint that the putative reward-prediction error responses are too short-lived to account for the kinds of learning that they direct (Redgrave *et al.*, 1999; Redgrave and Gurney, 2006). The basis for this challenge appears to relate to the issue of how a short burst of dopaminergic activity, and its putative translation into dopamine release in target areas, could account for physiological changes (like no dopaminergic response to a future expected reward) well into the future. To our understanding this is not a relevant challenge to RL models of dopamine function: in the models there is a clear distinction between the carriers of state value (which bridge the temporal gaps) and the reward-prediction errors (which are phasic). While the latter function is ascribed to dopamine, the former never was. Rather, it is presumed to be carried by cortical and/or striatal neurons that show sustained firing.

A related issue has emerged due to comparisons across vastly different experimental methods and time-scales. Longer-term changes in dopamine, as measured by a technique known as microdialysis, are marshaled as evidence for dopamine's role in what has been called *incentive salience*. Much has been written on this subject, but the argument is well summarized in Berridge (2007). Here, we would like to clarify what we think is limiting about the nature of the debate. In his summary of the issue, Kent Berridge states

> Debate continues over the precise causal contribution made by mesolimbic dopamine systems to reward. There are three competing explanatory categories: "liking", learning, and "wanting". Does dopamine mostly mediate the hedonic impact of reward ("liking")? Does it instead mediate learned predictions of future reward, prediction error teaching signals and stamp in associative links (learning)? Or does dopamine motivate the pursuit of rewards by attributing incentive salience to reward-related stimuli ("wanting")?
>
> (Berridge, 2007: 391).

In our view, the confusion here derives from at least three clear sources: (1) setting up the problem as though these separate questions are mutually exclusive; (2) comparing qualitative explanations in psychological terms like "wanting" and "liking" to quantitative models that match dopamine spike data to differential equations; and (3) a comparison of dopamine changes at vastly different time-scales. In Berridge's review, microdialysis measurements of dopamine levels under various behavioral challenges are compared to the reward-prediction error models that account for spike data. To be clear about the scope of the RL model of dopaminergic function, it applies strictly to rapid transients in spike rates in the 50–250 millisecond range and does not apply to other time-scales of dopaminergic modulation that may well carry other information important for cognitive processing and behavioral control. The examples above analyzed by Kakade and Dayan (2002) illustrate this issue. For example, the temporal-difference model is agnostic with regard to baseline dopamine levels or even fluctuations on slightly slower time-scales (e.g., minutes to hours). Consequently, the model would not account for microdialysis results whose measurements lie in these temporal regimes.

The last challenge is really an area of opportunity. Creatures ranging from sea slugs (e.g., *Aplysia*) to humans appear to be equipped with reinforcement learning systems – systems that broadcast some kind of evaluation signal to widespread regions of the nervous system and influence learning and decision making. A large body of neurobiological and behavioral data support such a position. In general, RL systems can be quite fast and efficient at learning, provided that the creature is pre-equipped with representations appropriate to the RL problems that it will face (see Dayan and Abbott, 2001). In the absence of appropriate representations, RL systems often perform miserably. Not much is known about how RL problems are represented in the human brain or how such representations should be modified by experience. This is an open future challenge for RL models; that is, to design experimental probes that reveal the structure of underlying representations.

CONCLUSIONS

To summarize, computational models of learning have done much to advance our understanding of decision making in the last couple of decades. Temporal-difference reinforcement learning models have provided a framework for optimal online model-free learning, which can be used by animals and humans interacting with the environment in order to learn to predict events in the future and to choose actions such as to bring about those events that are more desirable. Investigations into the decision-making behavior of both animals and humans support the existence of such a mechanism, controlling at least some types of rational behavior. The prediction-error hypothesis of dopamine has further linked these algorithmic ideas to possible underlying neural substrates, specifically, to learning and action selection in the basal ganglia modulated by phasic dopaminergic signals. Converging evidence from a wide variety of recording and imaging methods supports this hypothesis.

It seems that reinforcement learning has been most powerful (unfortunately for neuroscience, almost unique) in tying together the three levels – computation, algorithm, and implementation (Marr, 1982) – into one coherent framework that is used not only for gleaning understanding, but also for shaping the next generation of experimental investigations. Whether the theory can be elaborated to account for results of future experimentation without losing its simplicity and elegance, or whether it is eventually abandoned and replaced by a newer generation of computational learning theories, reinforcement learning has already left its permanent mark on the study of decision making in the brain.

Acknowledgments

This work was supported by the Human Frontiers Science Program (YN), The Kane Family Foundation (PRM), The Dana Foundation and Autism Speaks (PRM), National Institute on Drug Abuse (R01 DA11723 to PRM), National Institute of Neurological Disorders and Stroke (R01 NS045790 to PRM), and The Angel Williamson Imaging Center(PRM).

References

Baird, L.C. (1995). Residual algorithms: reinforcement learning with function approximation. In: A. Prieditis and S. Russell (eds), *Proceedings of the 12th International Conference on Machine Learning* (IMLL 95). San Mateo, CA: Morgan Kaufman, pp. 30–37.

Barto, A.G. (1995). Adaptive critic and the basal ganglia. In: J.C. Houk, J.L. Davis, and D.G. Beiser (eds), *Models of Information Processing in the Basal Ganglia*. Cambridge, MA: MIT Press, pp. 215–232.

Barto, A.G., Sutton, R.S., and Anderson, C.W. (1983). Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE Trans. Systems Man Cyber.* 13, 834–846.

Barto, A.G., Sutton, R.S., and Watkins, C.J.C.H. (1989). Sequential decision problems and neural networks. In: D.S. Touretzky (ed.), *Advances in Neural Information Processing Systems 2*. Cambride, MA: MIT Press, pp. 686–693.

Barto, A.G., Sutton, R.S., and Watkins, C.J.C.H. (1990). Learning and sequential decision making. In: M. Gabriel and J. Moore (eds), *Learning and Computational Neuroscience: Foundations of Adaptive Networks*. Cambridge, MA: MIT Press, pp. 539–602.

Bayer, H.M. and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141.

Bayer, H.M., Lau, B., and Glimcher, P.W. (2007). Statistics of midbrain dopamine neuron spike trains in the awake primate. *J. Neurophysiol.* 98, 1428–1439.

Behrens, T.E.J., Woolrich, M.W., Walton, M.E., and Rushworth, M.F.S. (2007). Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221.

Bellman, R.E. (1957). *Dynamic Programming*. Princeton, NJ: Princeton University Press.

Bergstrom, B.P. and Garris, P.A. (2003). "Passive stabilization" of striatal extracellular dopamine across the lesion spectrum encompassing the presymptomatic phase of Parkinson's disease: a voltammetric study in the 6-OHDA lesioned rat. *J. Neurochem.* 87, 1224–1236.

Berns, G.S., McClure, S.M., Pagnoni, G., and Montague, P.R. (2001). Predictability modulates human brain response to reward. *J. Neurosci.* 21, 2793–2798.

Berridge, K.C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacol. (Berl.)* 191, 391–431.

Berridge, K.C. and Robinson, T.E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Rev.* 28, 309–369.

Bertsekas, D.P. and Tsitsiklis, J.N. (1996). *Neuro-dynamic Programming*. London: Athena.

Bush, R.R. and Mosteller, F. (1951). A mathematical model for simple learning. *Psychol. Rev.* 58, 313–323.

Christoph, G.R., Leonzio, R.J., and Wilcox, K.S. (1986). Stimulation of the lateral habenula inhibits dopamine-containing neurons in the substantia nigra and ventral tegmental area of the rat. *J. Neurosci.* 6, 613–619.

Coizet, V., Dommett, E.J., Redgrave, P., and Overton, P.G. (2006). Nociceptive responses of midbrain dopaminergic neurones are modulated by the superior colliculus in the rat. *Neuroscience* 139, 1479–1493.

Daw, N.D. and Touretzky, D.S. (2000). Behavioral results suggest an average reward TD model of dopamine function. *Neurocomputing* 32, 679–684.

Daw, N.D. and Touretzky, D.S. (2002). Long-term reward prediction in TD models of the dopamine system. *Neural Computation* 14, 2567–2583.

Daw, N.D., Kakade, S., and Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Networks* 15, 603–616.

Daw, N.D., O'Doherty, J.P., Dayan, P. *et al*. (2006). Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.

Day, J.J., Roitman, M.F., Wightman, R.M., and Carelli, R.M. (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.* 10, 1020–1028.

Dayan, P. and Abbott, L.F. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: MIT Press.

Dayan, P. and Kakade, S. (2000). Explaining away in weight space. In: T. Leen, T. Dietterich, and V. Tresp (eds), *Advances in Neural Information Processing Systems*, Vol. 12. Cambridge, MA: MIT Press, pp. 24–30.

Dayan, P., Kakade, S., and Montague, P.R. (2000). Learning and selective attention. *Nat. Neurosci.* 3, 1218–1223.

Delgado, M.R., Locke, H.M., Stenger, V.A., and Fiez, J.A. (2003). Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cogn. Affect. Behav. Neurosci.* 3, 27–38.

Doya, K. (2000). Reinforcement learning in continuous time and space. *Neural Computation* 12, 219–245.

Doya, K. (2002). Metalearning and neuromodulation. *Neural Networks* 15, 495–506.

Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902.

Floresco, S.B., West, A.R., Ash, B. *et al*. (2003). Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nat. Neurosci.* 6, 968–973.

Franklin, K.B.J. and McCoy, S.N. (1979). Pimozide-induced extinction in rats: stimulus control of responding rules out motor deficit. *Pharmacol. Biochem. Behav.* 11, 71–75.

Geisler, S. and Zahm, D.S. (2005). Afferents of the ventral tegmental area in the rat-anatomical substratum for integrative functions. *J. Comp. Neurol.* 490, 270–294.

Goto, Y. and Grace, A.A. (2005). Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nat. Neurosci.* 8, 805–812.

Grace, A.A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience* 41, 1–24.

Hampton, A.N., Bossaerts, P., and O'Doherty, J.P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J. Neurosci.* 26, 8360–8367.

Hollerman, J.R. and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309.

Horvitz, J.C. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* 96, 651–656.

Houk, J.C., Adams, J.L., and Barto, A.G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: J.C. Houk, J.L. Davis, and D.G. Beiser (eds), *Models of Information Processing in the Basal Ganglia*. Cambridge, MA: MIT Press, pp. 249–270.

Howard, R.A. (1960). *Dynamic Programming and Markov Processes*. Cambridge, MA: MIT Press.

Ikemoto, S. and Panksepp, J. (1999). The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Res. Rev.* 31, 6–41.

Jensen, J., Smith, A.J., Willeit, M. *et al*. (2007). Separate brain regions code for salience vs valence during reward prediction in humans. *Hum. Brain Mapp.* 28, 294–302.

Joel, D. and Weiner, I. (1999). Striatal contention scheduling and the split circuit scheme of basal ganglia-thalamocortical circuitry: from anatomy to behaviour. In: R. Miller and J. Wickens (eds), *Conceptual Advances in Brain Research: Brain Dynamics and the Striatal Complex*. New York, NY: Harwood Academic Publishers, pp. 209–236.

Joel, D., Niv, Y., and Ruppin, E. (2002). Actor–Critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Networks* 15, 535–547.

Kacelnik, A. (1997). Normative and descriptive models of decision making: time discounting and risk sensitivity. In: G.R. Bock and G. Cardew (eds), *Characterizing Human Psychological Adaptations: Ciba Foundation Symposium 208*. Chichester: Wiley, pp. 51–70.

Kakade, S. and Dayan, P. (2002). Dopamine: generalization and bonuses. *Neural Networks* 15, 549–559.

Kamin, L.J. (1969). Predictability, surprise, attention, and conditioning. In: B.A. Campbell and R.M. Church (eds), *Punishment and Aversive Behavior*. New York, NY: Appleton Century Crofts, pp. 242–259.

Kehoe, E.J. (1977). Effects of serial compound stimuli on stimulus selection in classical conditioning of the rabbit nictitating membrane response. PhD thesis, university of Iowa.

Knutson, B. and Gibbs, S.E.B. (2007). Linking nucleus accumbens dopamine and blood oxygenation. *Psychopharmacol (Berl.)* 191, 813–822.

Knutson, B., Adams, C.M., Fong, G.W., and Hommer, D. (2001a). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* 21, RC159.

Knutson, B., Fong, G.W., Adams, C.M. *et al*. (2001b). Dissociation of reward anticipation and outcome with event-related fmri. *NeuroReport* 12, 3683–3687.

Knutson, B., Fong, G.W., Bennett, S.M. *et al*. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fmri. *NeuroImage* 18, 263–272.

Kobayashi, Y. and Okada, K.-I. (2007). Reward prediction error computation in the pedunculo-pontine tegmental nucleus neurons. *Ann. N.Y. Acad. Sci.* 1104, 310–323.

Konda, V.R. and Tsitsiklis, J.N. (2003). On Actor-Critic algorithms. *SIAM J. Control Optimization* 42, 1143–1166.

Konorski, J. (1948). *Conditioned Reflexes and Neuron Organization*. New York, NY: Cambridge University Press.

Kremer, E.F. (1978). The Rescorla-Wagner model: losses in associative strength in compound conditioned stimuli. *J. Exp. Psychol. Animal Behav. Proc.* 4, 22–36.

Lewicki, M.S. and Olshausen, B.A. (1999). A probabilistic framework for the adaptation and comparison of image codes. *J. Opt. Soc. Am. A* 16, 1587–1601.

Li, J., McClure, S.M., King-Casas, B., and Montague, P.R. (2006). Policy adjustment in a dynamic economic game. *PLoS ONE* 1, e103.

Ljungberg, T., Apicella, P., and Schultz, W. (1992). Responses of monkey dopaminergic neurons during learning of behavioral reactions. *J. Neurophysiol.* 67, 145–163.

Logothetis, N.K. (2003). The underpinnings of the BOLD functional magnetic resonance imaging signal. *J. Neurosci.* 23, 3963–3971.

Lohrenz, T., McCabe, K., Camerer, C.F., and Montague, P.R. (2007). Neural signature of fictive learning signals in a sequential investment task. *Proc. Nat. Acad. Sci. USA* 104, 9493–9498.

Marr, D. (1982). *Vision: A Computational Approach*. San Francisco, CA: Freeman & Co.

Matsumoto, M. and Hikosaka, O. (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* 447, 1111–1115.

McClure, S.M., Daw, N.D., and Montague, P.R. (2003a). A computational substrate for incentive salience. *Trends Neurosci.* 26, 423–428.

McClure, S.M., Berns, G.S., and Montague, P.R. (2003b). Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38, 339–346.

McClure, S.M., Li, J., Tomlin, D. *et al*. (2004). Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* 44, 379–387.

Menon, M., Jensen, J., Vitcu, I. *et al*. (2007). Temporal difference modeling of the blood-oxygen level dependent response during aversive conditioning in humans: effects of dopaminergic modulation. *Biol. Psych.* 62, 765–772.

Miller, R. and Wickens, J.R. (1991). Corticostriatal cell assemblies in selective attention and in representation of predictable and controllable events. *Concepts Neurosci.* 2, 65–95.

Mirenowicz, J. and Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379, 449–451.

Montague, P.R., Dayan, P., Nowlan, S.J. *et al*. (1993). Using aperiodic reinforcement for directed self-organization. In: C.L. Giles, S.J. Hanson, and J.D. Cowan (eds), *Advances in Neural Information Processing Systems*, Vol. 5. San Mateo, CA: Morgan Kaufmann, pp. 969–976.

Montague, P.R., Dayan, P., and Sejnowski, T.J. (1994). Foraging in an uncertain environments using predictive hebbian learning. In: Tesauro and J.D. Cowan (eds), *Advances in Neural Information Processing Systems*, Vol. 6. San Mateo, CA: Morgan Kaufmann, pp. 598–605.

Montague, P.R., Dayan, P., Person, C., and Sejnowski, T.J. (1995). Bee foraging in uncertain environments using predictive Hebbian learning. *Nature* 377, 725–728.

Montague, P.R., Dayan, P., and Sejnowski, T.J. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *J. Neurosci.* 16, 1936–1947.

Montague, P.R., McClure, S.M., Baldwin, P.R. *et al*. (2004). Dynamic gain control of dopamine delivery in freely moving animals. *J. Neurosci.* 24, 1754–1759.

Morris, G., Arkadir, D., and Nevet, A. (2004). Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* 43, 133–143.

Morris, G., Nevet, A., Arkadir, D. *et al*. (2006). Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* 9, 1057–1063.

Ng, A.Y., Harada, D., and Russell, S. (1999). Policy invariance under reward transformations: Theory and application to reward shaping. In: *Proceedings of the Sixteenth International Conference on Machine Learning*. San Francisco, CA: Morgan Kaufmann, pp. 278–287.

Nicola, S.M., Surmeier, J., and Malenka, R.C. (2000). Dopaminergic modulation of neuronal excitability in the striatum and nucleus accumbens. *Annu. Rev. Neurosci.* 23, 185–215.

Niv, Y. (2007a). Cost, benefit, tonic, phasic: what do response rates tell us about dopamine and motivation? *Ann. NY Acad. Sci.* 1104, 357–376.

Niv, Y. (2007b). The Effects of Motivation on Habitual Instrumental Behavior. Unpublished doctoral dissertation, The Hebrew University of Jerusalem.

Niv, Y., Daw, N.D., and Dayan, P. (2005a). How fast to work: response vigor, motivation and tonic dopamine. In: Y. Weiss, B. Schölkopf, and J. Platt (eds), *Advances in Neural Information Processing Systems*, Vol. 18. Cambridge, MA: MIT Press, pp. 1019–1026.

Niv, Y., Duff, M.O., and Dayan, P. (2005b). Dopamine, uncertainty and TD learning. *Behav. Brain Func.* 1, 6.

Niv, Y., Joel, D., and Dayan, P. (2006a). A normative perspective on motivation. *Trends Cogn. Science* 10, 375–381.

Niv, Y., Daw, N.D., and Dayan, P. (2006b). Choice values. *Nat. Neurosci.* 9, 987–988.

Niv, Y., Daw, N.D., Joel, D., and Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacol. (Berl.)* 191, 507–520.

Nomoto, K., Watanabe, T., and Sakagami, M. (2007). Dopamine responses to complex reward-predicting stimuli. *Soc. Neurosci. Abst.* 33, 749.5.

O'Doherty, J.P., Deichmann, R., Critchley, H.D., and Dolan, R.J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron* 33, 815–826.

O'Doherty, J., Dayan, P., Friston, K. *et al*. (2003). Temporal difference learning model accounts for responses in human ventral striatum and orbitofrontal cortex during Pavlovian appetitive learning. *Neuron* 38, 329–337.

O'Doherty, J.P., Dayan, P., Schultz, J. *et al*. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454.

Pagnoni, G., Zink, C.F., Montague, P.R., and Berns, G.S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nat. Neurosci.* 5, 97–98.

Pessiglione, M., Seymour, B., Flandin, G. *et al*. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442, 1042–1045.

Preuschoff, K., Bossaerts, P., and Quartz, S.R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51, 381–390.

Redgrave, P. and Gurney, K. (2006). The short-latency dopamine signal: a role in discovering novel actions? *Nat. Rev. Neurosci.* 7, 967–975.

Redgrave, P., Prescott, T.J., and Gurney, K. (1999). Is the short-latency dopamine response too short to signal reward error? *Trends Neurosci.* 22, 146–151.

Rescorla, R.A. (1970). Reduction in effectiveness of reinforcement after prior excitatory conditioning. *Learning Motiv.* 1, 372–381.

Rescorla, R.A. and Lolordo, V.M. (1968). Inhibition of avoidance behavior. *J. Comp. Physiol. Psychol.* 59, 406–412.

Rescorla, R.A. and Wagner, A.R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: A.H. Black and W.F. Prokasy (eds), *Classical Conditioning II: Current Research and Theory*. New York, NY: Appleton Century Crofts, pp. 64–99.

Reynolds, G.S. (1961). Attention in the pigeon. *J. Exp. Anal. Behav.* 4, 203–208.

Roesch, M.R., Calu, D.J., and Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nature Neurosci.* 10, 1615–1624.

Romo, R. and Schultz, W. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-intiated arm movements. *J. Neurophysiol.* 63, 592–606.

Salamone, J.D. and Correa, M. (2002). Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behav. Brain Res.* 137, 3–25.

Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340.

Samuel, A.L. (1959). Some studies in machine learning using the game of checkers. *IBM J. Res. Dev.* 3, 210–229.

Schönberg, T., Daw, N.D., Joel, D., and O'Doherty, J.P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *J. Neurosci.* 27, 12860–12867.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron* 36, 241–263.

Schultz, W., Apicella, P., Scarnati, E., and Ljungberg, T. (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. *J. Neurosci.* 12, 4595–4610.

Schultz, W., Apicella, P., and Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during succesive steps of learning a delayed response task. *J. Neurosci.* 13, 900–913.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Schwartz, A. (1993). Thinking locally to act globally: a novel approach to reinforcement learning. In: *Proceedings of the Fifth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 906–911.

Seymour, B., O'Doherty, J.P., Dayan, P. *et al*. (2004). Temporal difference models describe higher order learning in humans. *Nature* 429, 664–667.

Sutton, R.S. (1978). A Unified Theory of Expectation in Classical and Instrumental Conditioning. Unpublished Bsc thesis, Stanford University.

Sutton, R.S. (1988). Learning to predict by the method of temporal difference. *Machine Learning* 3, 9–44.

Sutton, R.S. and Barto, A.G. (1990). Time-derivative models of Pavlovian reinforcement. In: M. Gabriel and J. Moore (eds), *Learning and Computational Neuroscience: Foundations of Adaptive Networks*. Cambridge, MA: MIT Press, pp. 497–537.

Sutton, R.S. and Barto, A.G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Takikawa, Y., Kawagoe, R., and Hikosaka, O. (2004). A possible role of midbrain dopamine neurons in short- and long-term adaptation of saccades to position-reward mapping. *J. Neurophysiol.* 92, 2520–2529.

Tobler, P.N., Dickinson, A., and Schultz, W. (2003). Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J. Neurosci.* 23, 10402–10410.

Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642–1645.

Tom, S.M., Fox, C.R., Trepel, C., and Poldrack, R.A. (2007). The neural basis of loss aversion in decision-making under risk. *Science* 315, 515–518.

Ungless, M.A., Magill, P.J., and Bolam, J.P. (2004). Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science* 303, 2040–2042.

Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412, 43–48.

Watkins, C.J.C.H. (1989). Learning with Delayed Rewards. Unpublished doctoral dissertation, Cambridge University, Cambridge.

Weiner, I. and Joel, D. (2002). Dopamine in schizophrenia: dysfunctional information processing in basal ganglia-thalamocortical split circuits. In: G.D. Chiara (ed.), *Handbook of Experimental Pharmacology* Vol. 154/II, *Dopamine in the CNS II*. Berlin: Springer Verlag, pp. 417–472.

Werbos, P.J. (1977). Advanced forecasting methods for global crisis warning and models of intelligence. *General Systems Yearbook* 22, 25–38.

Wickens, J. and Kötter, R. (1995). Cellular models of reinforcement. In: J.C. Houk, J.L. Davis, and D.G. Beiser (eds), *Models of Information Processing in the Basal Ganglia*. Cambridge, MA: MIT Press, pp. 187–214.

Williams, R.J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning* 8, 229–256.

Willner, P., Towell, A., and Muscat, R. (1987). Effects of amphetamine and pimozide on reinforcement and motor parameters in variable-interval performance. *J. Psychopharmacol.* 1, 140–153.

Wise, R.A. (1982). Neuroleptics and operant behavior: the anhedonia hypothesis. *Behav. Brain Sci.* 5, 39–53.

Wise, R.A. (2004). Dopamine, learning and motivation. *Nat. Rev. Neurosci.* 5, 483–495.

Wise, R.A., Spindler, J., de Wit, H., and Gerberg, G.J. (1978a). Neuroleptic-induced "anhedonia" in rats: pimozide blocks reward quality of food. *Science* 201, 262–264.

Wise, R.A., Spindler, J., and Legault, L. (1978b). Major attenuation of food reward with performance-sparing doses of pimozide in the rat. *Can. J. Psychol.* 32, 77–85.

Yu, A.J. and Dayan, P. (2002). Acetylcholine in cortical inference. *Neural Networks* 15, 719–730.

Yu, A.J. and Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron* 46, 681–692.

Zald, D.H., Boileau, I., El-Dearedy, W. *et al*. (2004). Dopamine transmission in the human striatum during monetary reward tasks. *J. Neurosci.* 24, 4105–4112.

CHAPTER

23

The Neurobiological Foundations of Valuation in Human Decision Making Under Uncertainty

Peter Bossaerts, Kerstin Preuschoff and Ming Hsu

OUTLINE

INTRODUCTION

The goal of this chapter is to review recent neurobiological evidence to improve our understanding of human valuation under uncertainty. Although ultimately interested in human behavior, we will borrow from studies of animals with related brain structures, namely, non-human primates. Specifically, we wish to explore how valuation is accomplished. As we shall see, the evidence rejects a pure "retrieval from memory" model; instead, values are *computed*. This raises the issue: what computational model(s) are being used? Since actual choice can be summarized in terms of a single-dimensional utility index as in expected utility or prospect theory, we want to know how such an index is computed, and to understand the effect of perceptual biases on this computation, as well as the role of emotions. How does the computational model generate the risk aversion that we may see in choices? Or, in ambiguous situations, how is ambiguity aversion revealed in choices (Hsu *et al.*, 2005; Huettel, *et al*. 2006; Bali *et al.*, 2008) and what model underlies it – for example, alpha-maxmin preferences (Ghirardato *et al.*, 2004), anticipated regret (Segal, 1987), or some other?

A natural question for economists is, if choice can be represented "as if" some utility index is maximized, why should we bother studying the computational aspects behind choice? Our quest for the computational model underlying valuation has two purposes. First, we hope to be able to improve choice prediction. We consider the utility index model as a reduced-form model of choice, which, in analogy with reduced-form models of the macro-economy, is sensitive to circumstances – i.e., its parameters need to be revised every time one moves from one framework to another. This is the so-called *Lucas critique* (Lucas, 1976). This sensitivity, or lack of robustness, has been frequently noted in the literature as "framing effects." Our computational model, in contrast, is meant to be structural, and hence has a chance of outperforming the reduced-form utility index model. Of course, the proof is in the pudding …

A computational model should not only enable us to accommodate framing biases; it should also be capable of capturing part, if not all, of the randomness one usually adds to a utility-based choice model to accommodate the erratic nature of observed choice. The traditional approach in applied economics has been to consider the utility index itself to be random, and to attribute certain distributional properties to this randomness that ultimately leads to a choice model based on logistic regression (McFadden, 1974). The goal of exploring computational models of choice is to provide a different (computational) basis for the erratic nature of observed choice.

Artificial intelligence is the second reason why we are interested in the actual neurobiological computational model behind observed choices. Modern decision theory has been quite successful in prescribing learning and choice in well-defined situations. Once the complexity of the problem increases, however, decision theory has little concrete to propose. One case in point is the Iowa Gambling Task (IGT) (Bechara and Damasio, 2005), which is a complex four-armed bandit problem. To date, no-one has been able to spell out the optimal strategy in this task. Humans, however, can quickly learn the ordering (in terms of reward and risk) of the four bandits (Bechara *et al*., 1997).

However, the learning model behind human choice in difficult situations such as the IGT is almost surely not Bayesian, despite the popularity of this approach. Open-ended situations, almost by definition, involve high-dimensional parameter spaces, and it is well known that Bayesian learning generically fails to converge when the parameter space becomes large, while simple learning algorithms based on classical statistics often do a very good job in capturing the essence of the stochasticity at hand (Diaconis and Freedman, 1986). This in turn leaves us the question, how do humans learn in such situations? What computational model are they using? If we knew this model, perhaps we could enrich decision theory to make it work in more open-ended or even ill-defined decision situations.

One clue in the search is provided by the plethora of brain lesion data, much of it using the IGT (despite problems with interpretability – see Dunn *et al*., 2006), supporting the idea that specific brain regions underpin these computations. Of particular interest are the ventromedial prefrontal cortex and the amygdala, as patients with lesions to these brain regions are impaired in the IGT while at the same time showing normal or even superior performance in working memory, IQ, and decisions and games that are less open-ended (Bechara *et al*., 1997, 1999).

In this chapter, we focus on valuation in the context of choice under uncertainty. The computational model that emerges from a study of the neurobiological data is one where expected rewards are traded off against risk (or, in the case of risk-seeking behavior, risk is added to expected reward). This model has been popular in finance, where risk is usually represented in terms of reward variance (although more advanced models take into account higher moments) – hence its name: the *mean-variance model*.

The link between the utility models representing actual choice under uncertainty, such as expected utility and prospect theory, on the one hand, and the mean-risk model, on the other, is often clarified by means of Taylor series expansions (Bali *et al*., 2008; see also Box 23.1). One of the goals of this chapter is to demonstrate that the logic of a computational model based on a tradeoff between expected reward and risk can be extended to choice under ambiguity as well.

Throughout this chapter, we take the position that valuation stands on its own and makes sense even in the absence of choice. It makes sense because computing values takes energy, and hence an organism that has learned to correctly compute the values of options it is forced to take on (imperative trials) will also be better able to determine the adequate choices when there is freedom (free-choice trials). As we shall see, brain activation when a subject is faced with gambles with uncertain outcomes is comparable across imperative and free-choice trials.

When valuation is a pursuit distinct from choice, values revealed through choice may be different from the valuations that emerge from the (choice-independent) computations. So are economists interested in knowing what the latter are? After all, economists are only interested in choice. We shall spend some time elaborating on this issue. Suffice it to say here that the distinction is important for welfare assessment.

The remainder of this chapter is organized as follows. The following section elaborates on the distinction between values the brain computes and the values revealed through choice. We then review the neurobiological foundations of valuation under pure

BOX 23.1

MEAN-VARIANCE UTILITY

The expected utility approach, in which the decision maker maximizes a weighted sum of the utilities, dominates in economics. In finance, however, mean-variance (or risk-return) models provide the central approach to decision making. In it, the decision maker trades off between mean (return) and variance (risk). The two approaches are related. There are conditions under which the two models are identical, such as when returns are approximately normal, or if the utility function is quadratic (Markowitz, 1952). More generally, the two can be linked by assuming that the mean-variance approach is an approximation of the utility function through a Taylor series approximation.

Specifically, given a utility function $U(R)$, where R is the rate of return in the current period, we can implement a second-order Taylor series approximation around the mean $\mu = E(\mathrm{R})$, such that

$$U(R) \approx U(\mu) + U'(\mu)(R - \mu) + U''(R - \mu)^2.$$

The expected utility is therefore

$$E(U(R)) \approx U(\mu) + U''(\mu)\sigma^2,$$

where the second term $\sigma^2 = Var(R)$ is the variance of the return.

For more on the relationship between these approches, see Chapter 10.

risk. The final section extends the logic of expected reward and risk to situations with ambiguity. We will not explore the neurobiological foundations of the translation of valuation into choice, as these are covered in later chapters in this volume.

DISTINGUISHING VALUATION AND CHOICE

In many human endeavors, valuation is performed even in the absence of any immediate necessity to choose. Again, finance is a case in point, in part because financial valuation is often complex and time-consuming, while good choice opportunities are rare and short-lived. The cost of computing values provides a normative rationale for why valuation may be done in the absence of free choice. Figure 23.1 provides strong evidence that valuation is performed even in imperative trials, and in a way that is relevant for free-choice trials. It shows how well a valuation model fit to neuronal activation during imperative trials can predict choice in free-choice trials. In fact, the neurobiological choice model predicts choices better than a utility-index based model estimated from the choices themselves! This demonstrates not only that valuation is done during imperative trials, but that the resulting values are relevant for choice in free-choice trials as well.

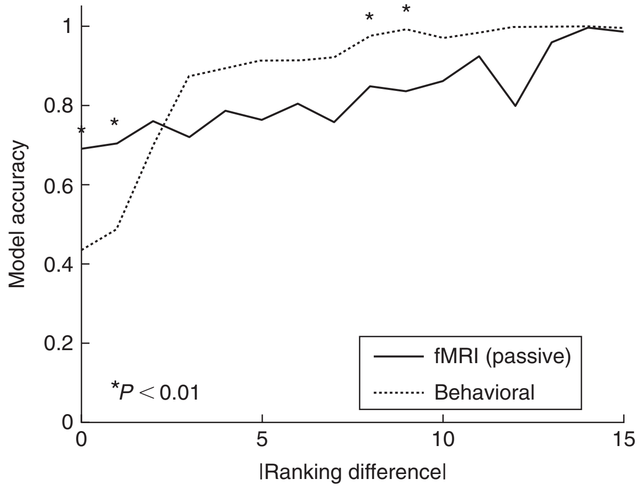


FIGURE 23.1 Out-of-sample prediction accuracy (percentage correct) of two models of binary choice, as a function of the ranking difference of the two options revealed through choice. The fMRI (passive) model uses as inputs activation in brain regions identified to be involved in valuation during purely imperative trials. The behavioral model is based on rankings revealed in choice in free-choice trials. The parameters of both models are fit on free-choice trials. Prediction accuracy is out-of-sample, i.e., it is measured in free-choice trials not used for fitting the models. Neural activation identified in imperative trials (the fMRI model) predicts choices better than actual choice in free-choice trials when the two available options are close in ranking. Adapted from Berns *et al.* (2007).

Although brain activation during imperative trials reflects valuations that are compatible with the values revealed in free-choice trials, and hence that brain activation in imperative trials can be used to predict choice in free-choice trials (Figure 23.1), the fit is not 100%. These neurobiological data suggest that there are (at least) two value signals: one revealed through activation in brain regions not directly involved in the physical implementation of choice, and a second one revealed through activation of the neurons controlling the physical act of choice. Given economists' interest in choice, to what extent might they be interested in knowing the former value signal?

The most important argument concerns welfare. What is it that the subject really wanted? Indeed, the mere presence of two (or more) valuation signals suggests that there may be a difference between what subjects wanted and what they did. We may never be able to determine what exactly subjects wanted (i.e., to extract their true preferences), but a better understanding of how the various value signals are put together may help us to better appreciate the difference between revealed and true preferences.

It is fundamental to question coincidence of revealed and true preferences (psychologists would distinguish between decision and experienced utility – see Kahneman *et al*., 1997). Is a utility-index model such as prospect theory a preference theory? Prospect theory is known to fit choice well, but do the values of its utility index reflect real preferences? The insistence on deriving choice-theoretic axiomatic foundations for prospect theory (Kahneman and Tversky, 1992) indicates that economists consider prospect theory to be a preference theory. However, the risk-seeking attitude towards losses that is implied by prospect theory, for instance, may not reflect context-adapted goal-oriented behavior, but a "default action" that is appropriate to maximize experienced utility only on average. The default action is robust to lapses of attention, and hence is optimal for an organism that has to spend effort to interpret the true meaning of stimuli around itself. It may be considered a "habit," but the neuroscience literature reserves the term *habit* for actions that are stimulus-sensitive but goal-insensitive (Yin *et al*., 2004). In contrast, default actions are stimulus-insensitive and goal-oriented. They are not instincts, either, because they can be learned.

Recent neurobiological evidence provides support for the notion of default actions. Choice in loss situations appears to be associated with activation in amygdala, irrespective of a subject's tendency to take on risk. Orbitofrontal cortex (OFC) activation, however, decreases with subjects' tendency to become risk-seeking with losses (Figure 23.2). These data suggest a very different valuation model underlying choice, based on default actions that need to be overcome in situations where these actions are inappropriate – yet not everyone manages to over-ride the default. In this interpretation of the data, amygdala activation engages the default action, while OFC is engaged in over-riding this choice when inappropriate. Perhaps no one wants to be risk-seeking for monetary losses,

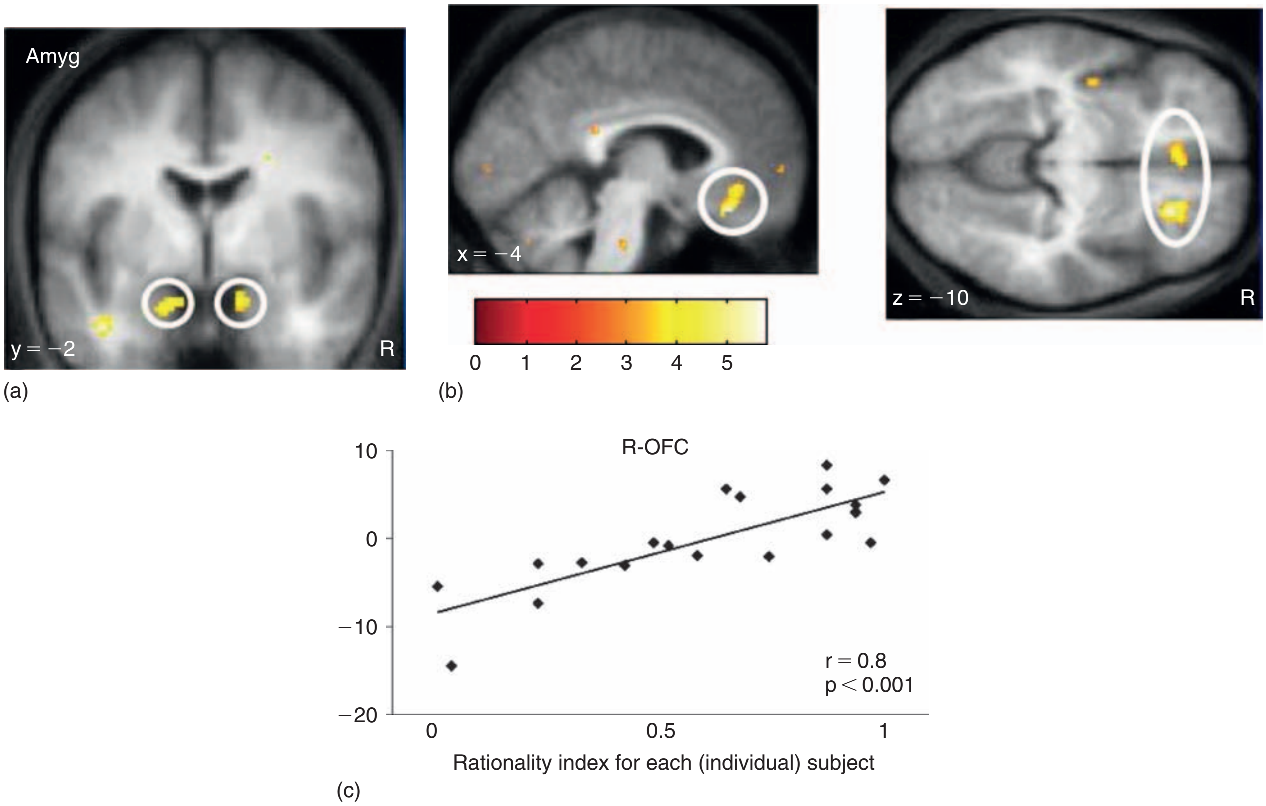


FIGURE 23.2 (a) Amygdala activation correlates with contrast between lotteries cast in terms of gains and in terms of losses. (b, c) Contrast in prefrontal cortex activation between lotteries cast in terms of gains and in terms of losses increases with subjects' rationality index. This index is measured as the difference between the proportion of trials in which subjects chose the gamble in the loss frame, as compared to the gain frame, and then linearly transformed such that 1 is the most rational. Adapted from De Martino *et al*. (2006).

but many subjects cannot overcome their tendency to become risk-seeking under adverse conditions.

More generally, it is not so much the value signals that may be discerned in, for example, prefrontal cortex (vmPFC) which are of importance. More relevant is the way the signals are computed: which brain regions are involved? What is the nature of their signals? How do they correlate with behavior?

The potential difference between true and revealed preferences is, in principle, not important for predicting choice, yet it is of utmost important for welfare. At a mundane level, it should be asked whether bankers and brokers ought to engage in eliciting the right prospect-theoretic utility index from their customers through, for example, questionnaires on past or hypothetical choices, to be used to determine optimal portfolios. This only makes sense if prospect theory captures true preferences. If it does not, what is the point of deriving optimal portfolio choice from prospect-theoretic preference profiles (Gomes, 2005)?

In fact, the mere difficulty of computing optimal choices should cast doubt on the interpretation of prospect theory as a preference theory. Indeed, an organism that truly behaves in a prospect-theoretic way must be endowed with tremendous capacity to compute optimal choices. To put it more bluntly, perhaps organisms would be better off with logarithmic preferences, for which choices are easily computed – for example, one can be myopic (Hakansson, 1971). In addition, this has the advantage that choices maximize survival probabilities (Blume and Easley, 1992)!

Data in De Martino *et al.* (2006) could be re-interpreted to mean that choice is the result of valuing the appropriateness of some default action, given the stimuli at hand, and, if inappropriate, effort to overcome one's tendency to implement the default action. Recent single-unit recording of neurons in the monkey brain indicates that caudate plays a crucial role in the valuation of default actions. Figure 23.3 provides evidence.

In the task that generated Figure 23.3, a monkey had to fixate on the center of the screen and at a certain time move its eyes (execute a "saccade") in the direction indicated by a prior stimulus. Only one direction was rewarded, however. To be precise, the chance that one direction would be rewarded was an order of magnitude higher than for the other direction. As it turns out, certain caudate neurons increase their firing rate at trial onset, even before the stimulus, and hence the required direction of motion of the eyes, as well as the amount of the reward, is known. This firing reflects valuation of the default action – namely, to move the eyes in the direction that is most likely to be rewarded. When the stimulus appears, neuronal firing either increases, reflecting confirmation that the default action is correct, or decreases, signaling that the default action is inappropriate (Figure 23.3). In fact, if the percentage of correct saccades is an indication, effort is involved in avoiding the default action. Indeed, Lauwereyns *et al.* (2002) reported that the monkey made more mistakes (i.e., moved its eyes in the wrong direction) when the stimulus required a saccade in the less rewarding direction.

The firing pattern of the caudate neurons also casts doubt on random-utility models (McFadden, 1974) as an explanation of behavior that is at odds with maximization of a fixed utility index. The pattern suggests that the apparent "erratic" behavior of the monkey (its mistakes) is the result of its inability to overcome the default action – for example, if the monkey moved its eyes in the direction that was rewarded with higher probability while the stimulus instructed it otherwise, it was not that its risk tolerance had increased so much that it wanted to take the chance that its visual

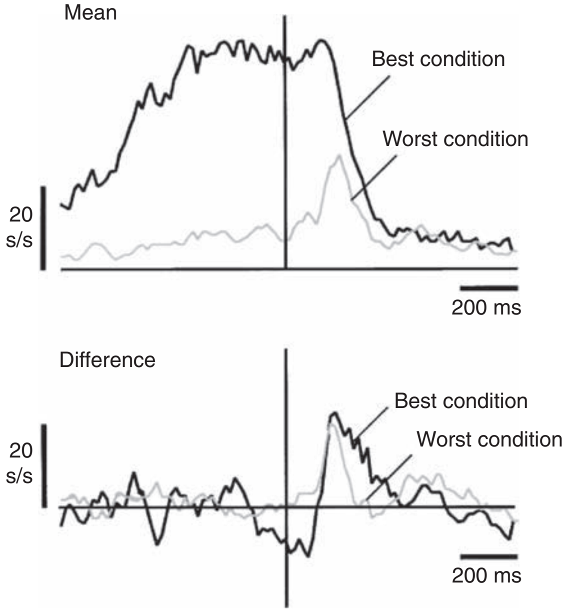


FIGURE 23.3 Average firing rates of a cluster of caudate neurons in the monkey brain that encode the value of eye movements (saccades) in one "preferred" direction. In trials where this direction generates higher reward on average ("Best condition"), increase in firing starts several hundred milliseconds (ms) before stimulus emerges (vertical line) that indicates which direction will be rewarded (top graph, labeled "mean"). Anticipatory increases in firing rates are absent in trials where the "preferred" direction receives the lower reward on average ("Worst condition"). After stimulus presentation, an increase is recorded in average firing rates across trials where the stimulus indicates reward for a saccade in the "preferred" direction, relative to average firing rates when saccades are instructed in the opposite direction. The difference (bottom graph) is the same for both "Best condition" and "Worst condition" trials. From Lauwereyns *et al.* (2002).

BOX 23.2

RANDOM-UTILITY MODELS

Random utility models form a class of statistical models widely used to describe actually observed choice behavior. Following standard revealed-preference principles, individual choices are assumed to maximize some underlying utility. The latter is, however, imperfectly observable. Utility, therefore, is split into a modeled part (a function of observable variables, common to all individuals up to a parameter vector) and an orthogonal, random error term (often referred to as "unobserved heterogeneity").

Specifically, the utility for alternative a is written as

$$U_a = V_a + \varepsilon_a,$$

where V_a is the modeled part of the utility, and ε_a the error term, capturing the uncertainty over the true U_a. The probability of observing alternative a being chosen in the choice set C is therefore

$$\Pr_C(Y = a) = \Pr(U_a = \max_{c \in C} U_c)$$

The uncertainty ε_a is frequently modeled as a normal or logistic distribution, yielding probit or logit models, respectively.

inference (of the instructions revealed by the stimulus) was wrong. The mistake was caused by the monkey's inability to overcome its default action (a saccade in the direction of reward with higher probability). Conversely, if the monkey did not make a mistake and moved its eyes in the direction of lower probability of reward, as instructed through the stimulus, it was not that the monkey became less risk-tolerant and did not want to take chances; instead, it overcame its natural inclination to saccade in the direction that tended to be rewarded with higher probability.

Even if the random-utility model (Box 23.2) and the default-action valuation model predict choice equally well, economists should be aware that neurobiological evidence favors the latter. This is because the welfare implications of the two models are diametrically opposed. According to the random utility model, choice is always optimal. The agent makes different choices in two instances of the same situation only because utility fluctuates. In the default-action valuation model, however, choice is often sub-optimal – for example, the monkey wanted to move its eyes in the opposite direction, but could not because of the effort required to overcome its habit. Now imagine that some external correcting device can be provided using an effort that is less than that required for the monkey personally to overcome its action bias. In that case, welfare will improve. See Bernheim and Rangel (2008) for further discussion of welfare assessment in a context where one needs to distinguish between "true" and "decision" utilities.

VALUATION UNDER PURE RISK: TRADING OFF RISK AGAINST REWARD

Pure risk is a situation where probabilities are known, either because they are given or because the agent has gone through lengthy training. Economists model choice under pure risk as if the agent maximizes a utility index based on a separation of probabilities (of all possible states of nature) and utilities (of rewards in a state). This separation is known as *probabilistic sophistication* (Marinacci, 2002), and is common to virtually all preference models under pure risk (including prospect theory).

There appears, however, to be little neurobiological evidence for value computation based on separation of (state) probabilities and (reward) utilities. Instead, brain activation data suggest a separation of encoding of expected reward and of risk.

Evidence of encoding of expected reward, and the crucial role of the dopaminergic system, goes back a long way, to the investigation of firing of dopaminergic neurons in the monkey midbrain in the face of stochastic rewards (Hollerman and Schultz, 1998). Closer study of the activation of these neurons has revealed that the signals actually reflect *errors* of predicting reward, and these errors form the crucial component of a simple yet versatile and powerful algorithm to learn expected rewards based on the Rescorla-Wagner reinforcement learning rule and extendible to complex multiple stimuli-reward situations (Montague *et al.*, 1996; McClure *et al.*, 2003; O'Doherty *et al.*, 2003). The prediction error equals the difference between actual or updated and (previously) expected (sums of discounted) rewards. A number of chapters in this volume provide details, so we shall not elaborate here (see Chapters 21, 22, 24, and 26).

One crucial element is missing in this account of (expected) reward learning, namely, risk. Specifically, to learn at the optimal rate, it is important that the decision maker assesses the risk of making a prediction error. The reasoning is simple. If expecting to make large prediction errors (i.e., if the risk is expected to be high), then one should not change one's prediction much based on a large actual prediction error. In other words, the learning rate should be low[1].

There are (at least) two ways in which risk assessment could affect learning. One is directly through the learning rate; another is through adjustment of the prediction error. Indeed, a simple scaling of prediction errors by their expected size accomplishes the same (see Preuschoff and Bossaerts, 2007, for details). Evidence for such a type of "adaptive encoding" was recently discovered in firing of dopamine neurons in the monkey brain. Figure 23.4 shows the average firing rates and single-trial firings of dopamine neurons for three types of trials, in the order (top to bottom) of increasing expected size of the prediction error. When reward occurs, the prediction errors are insensitive to the level of anticipated risk, demonstrating that they are somehow "scaled." This finding is further explored in Chapter 21.

(Expected) reward learning is complex. Not only will it be affected by the expected size of the prediction error (risk); other considerations also play a role, such as how much an optimal predictor might co-vary with the prediction error. For instance, in situations where the environment changes rapidly, past prediction errors become obsolete fast, and hence prediction should rely more on recent prediction errors. Effectively, this means that the learning rate should increase. This intuition has a rigorous underpinning (see Preuschoff and Bossaerts, 2007). Behrens *et al.* (2007) recently provided evidence that humans do adjust their learning rate to the degree of stability.

In addition to being necessary for optimal learning, risk-encoding could also play a role in the decision maker's risk *attitude*. Risk-encoding is actually *presumed* to occur in risk-sensitive organisms. How else would their behavior be affected by risk? Not surprisingly, the evidence of risk-encoding in the human and non-human primate brain is overwhelming. Regions where activation appears to be sensitive to risk include the insula, anterior cingulate cortex (ACC) and inferior frontal gyrus (IFG) (see, for example, Critchley *et al.*, 2001; Paulus *et al.*, 2003; Huettel *et al.*, 2005). Some of these regions, such as insula, seem to encode risk *exclusively* (Preuschoff *et al.*, 2006, 2008).

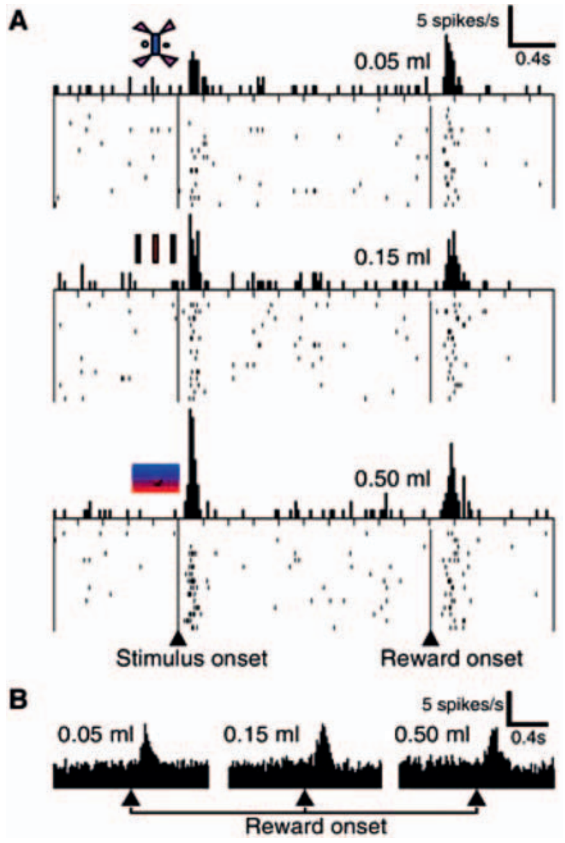


FIGURE 23.4 Single-trial firing rates and corresponding histogram for a single dopamine neuron in ventral tegmental area (a) and histogram of group firing rates (b) in imperative trials where a monkey is presented with stimuli that indicate the size of a random (50% chance) juice reward; results are shown only for rewarded trials. Firing at stimulus onset reflects the size of the expected juice reward. Firing at reward onset reflects the prediction error (positive, since only rewarded trials are retained for the analysis). The firing at reward delivery does not, however, reflect the size of the prediction error, which increases with the size of the reward (e.g., 0.50 ml against an expectation of 0.25 ml in the third case, or 0.05 ml against an expectation of 0.025 ml in the first case). Firing at reward delivery reflects a scaled prediction error: the prediction error divided by the anticipated size of the prediction error (prediction risk). From Tobler *et al.* (2005).

Moreover, in simple monetary gambles, activation seems to reflect *variance* (or its square root, standard deviation), i.e., the expectation of the squared prediction errors (Preuschoff *et al.* 2006, 2008). As such, the encoding provides a key parameter with which to compute the utility of a risk-averse agent. Indeed, a Taylor series expansion of an expected utility index demonstrates that risk is to be measured in terms of reward variance; all other aspects of risk (skewness, kurtosis, etc.) are secondary. For discussion of applications to finance, see Bali *et al.* (2008), for neuroimaging data on possible encoding of skewness and kurtosis, see Hsu (2006).

As with expected reward signals in the dopaminergic system, activation correlating with risk in some regions actually reflects risk-prediction *errors* – i.e., the difference between the square-size of the prediction

---

[1] A few neuroscientists have started to incorporate risk in reinforcement learning, referring to the (prediction) risk as *expected uncertainty* (Yu and Dayan, 2003). For a rigorous treatment, see Preuschoff and Bossaerts, 2007.

error and its expectation (the variance). Specifically, phasic activation in anterior insula exhibits strong correlation with risk-prediction errors (Preuschoff *et al.*, 2008). Little is known, however, about the precise neurobiological architecture and algorithms of risk learning. We only know that risk prediction and risk-prediction errors are encoded.

One may object that a risk-prediction error can be recorded in the paradigm of Preuschoff *et al*. (2008). Indeed, the risk-prediction error in a trial is irrelevant to making better forecasts in future trials, as the trials are independent. That is, there is no relevant learning. There are three reactions to this objection. First, this finding is not unique to risk-prediction errors. Reward-prediction errors can similarly be recorded in situations where learning is irrelevant because trials are independent (see, for example, Fiorillo *et al.*, 2003; Preuschoff *et al.*, 2006). Second, encoding of prediction errors is relevant for cases where learning unexpectedly *does* matter. For example, if an experimenter tells a subject that trials are independent (as in Preuschoff *et al.*, 2006) yet is lying (unlike in Preuschoff *et al.*, 2006), then if the brain has prediction errors at hand it can manage the situation better[2]. In this sense, the emergence of prediction errors in situations where learning is irrelevant in principle is equivalent to the presence of valuation signals in imperative trials.

Third, prediction-error encoding does sometimes have behavioral implications. Choices across independent trials sometimes reflect belief updating as if outcomes are actually dependent. Such belief updating presumes encoding of prediction errors. A classical example is underweighting of small probability events in experiments with trial-by-trial reward feedback (see Hertwig *et al.*, 2004).

Finally, risk-encoding may play a role in learning, in addition to guiding choice. For learning to be effective, however, risk-encoding should be *objective*. In particular, risk aversion or risk tolerance should not affect risk perception. Imagine that more risk-averse agents overestimate risk. Since perceived risk decreases the learning rate, this would mean that risk-averse agents learned more slowly. That would give them a handicap in competition with agents who are more risk tolerant. To date, it is not known to what extent learning is affected by risk aversion, or to what extent the brain encodes a dual-risk signal – one subjective, to guide choice, and one objective, to drive learning.

The picture that emerges from the extant findings is one where the brain values risky gambles by evaluating their expected reward and risk separately (Figure 23.5). The separate evaluations are then merged to generate a total valuation signal, detectable in (for example) prefrontal cortex (PFC). Figure 23.6 shows how PFC activation increases with expected reward for all subjects, but decreases with risk for risk-averse subjects and increases with risk for risk-seeking subjects.

To date, it is not known how the various signals related to risk and expected reward in disparate brain regions are integrated to deliver the aforementioned total value signal in PFC. A recent study, however, used canonical correlations analysis to extract a value index from joint activation of insula, ventral striatum, and putamen, exploiting the recorded correlation in baseline activations across these regions. The resulting "neuronal" utility index was found to correlate significantly with activation in the very part of PFC identified in Tobler *et al*. (2007; see also Bruguier *et al.*, 2008).

EXTENDING THE REWARD–RISK COMPUTATIONAL MODEL TO AMBIGUITY

So far, we have considered only pure risk – i.e., situations where probabilities are known. The evidence points towards a computational model whereby the brain computes value by separately encoding expected reward and risk, and combining the results. Such a computational model is known to approximate well to many types of utility functions (Bali *et al.*, 2008), including prospect theory (Agren, 2006).

In many situations in real life, probabilities are unknown or only partially known. This is called *ambiguity* in decision theory (Ellsberg, 1961). A stylized fact in decision experiments is that many individuals are ambiguity averse (Camerer and Weber, 1992). A number of models exist that account for decisions under ambiguity. Perhaps the best known are the maxmin utility model of Gilboa and Schmeidler (1989), and its extension, the α-maxmin utility model (Ghirardato *et al.*, 2004). In the latter, the worst- and best-case scenarios provide anchor points for beliefs. In one extreme case ($\alpha = 1$), beliefs are entirely determined by the worst-case scenario, and the α-maxmin utility model coincides with that of Gilboa and Schmeidler (1989). In the other extreme case ($\alpha = 0$), beliefs are solely based on the best-case scenario. For $\alpha > 0.5$, the decision maker is ambiguity averse; for $\alpha < 0.5$, she is ambiguity seeking; at $\alpha = 0.5$, she is ambiguity neutral – i.e. her beliefs are the average of the worst-case and best-case scenarios.

---

[2] In fact, this is one of the prime reasons why deception is frowned upon in economic experiments, but it would not be a compelling one if people always blindly believed the experiment.

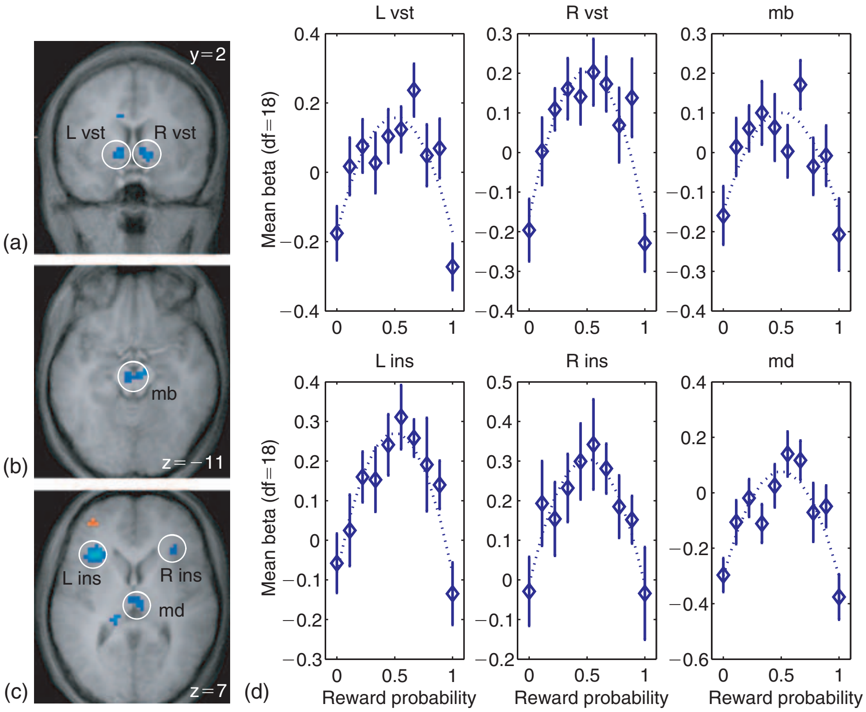


FIGURE 23.5 (Immediate) activation in striatal areas (a) increases with the probability of a fixed reward in the human brain in gambles where the probability of reward fluctuates between 0 and 1. As such, activation correlates with expected reward. (Delayed) activations in sub-cortical areas (b) and insula (c) change quadratically with probability of reward (d), with a peak at 50% chance of winning, and minima at certainty of loss or gain. As such, activation correlates with risk, measured as reward variance. Vertical line segments indicate 95% confidence intervals. From Preuschoff *et al.* (2006), and unpublished data.

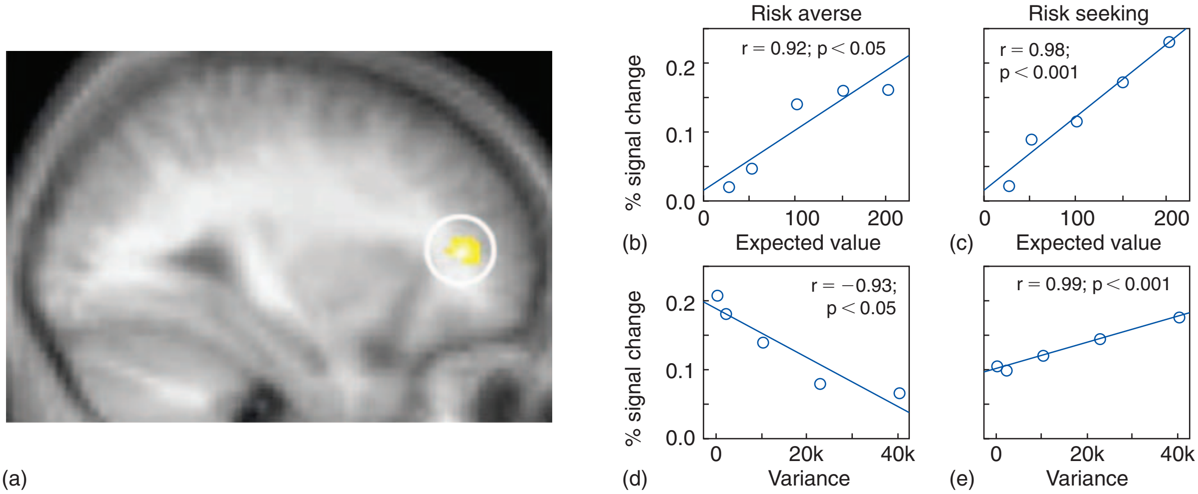


FIGURE 23.6 Activation in medial prefrontal cortex correlates with expected utility of a gamble in imperative trials (a). This activation increases in expected reward both for risk-averse (b) and risk-seeking (c) subjects; risk attitudes are measured through revealed preference in free-choice trials. The same activation decreases with risk of the gamble (measured as reward variance) for risk-averse subjects (d). It increases with risk for risk-seeking subjects (e). Adapted from Tobler *et al.* (2007).

For instance, the decision maker may bet on the drawing of red, green, and blue balls from an urn. She earns \$1 if the ball drawn is red or green; if the ball is blue, she loses \$1. One-third of the balls are red. The number of green and blue balls is unknown, but the decision maker knows that at least one-sixth of the balls are green, and one-twelfth are blue. A decision maker with $\alpha = 0$ will only consider the worst-case scenario, the minimum probability of winning, $p_{min} = 1/2$ $(=(1/3) + (1/6))$; this will fix her beliefs. When $\alpha = 1$, the maximum probability, p_{max}, determines her beliefs; $p_{max} = 11/12$ $(=1 - 1/12)$. In general, the belief that the decision maker wins equals $\alpha\ p_{min} + (1 - \alpha)\ p_{max}$. Without loss of generality, we can set the utility of winning \$1 equal to 1 and the utility of losing \$1 equal to 0. Hence, the general expression of the expected utility of the decision maker is $\alpha\ p_{min} + (1 - \alpha)\ p_{max}$. See Box 23.3 for further discussion of the α-maxmin model.

As with expected utility under pure risk, the α-maxmin utility model is equivalent (this time, it is *not* an approximation) to one where utility is computed based on a tradeoff between mean and risk. In this case, the mean is just the average of the minimum and maximum probabilities:

$$\bar{p} = \frac{1}{2} p_{min} + \frac{1}{2} p_{max}.$$

The risk c, in contrast, is determined by the amount ambiguity:

$$c = p_{max} - p_{min}.$$

Utility, then, is some weighted average of β and c:

$$U = \gamma_1 \bar{p} + \gamma_2 c.$$

Simple algebra reveals that:

$$\alpha = \frac{1}{2}\gamma_1 - \gamma_2.$$

$$1 - \alpha = \frac{1}{2}\gamma_1 + \gamma_2.$$

So $\gamma_1 = 1$ and $\gamma_2 = 1/2 - \alpha$.

Therefore, under both pure risk and ambiguity, there exists an equivalent representation of utilities in terms of a model that trades off mean against risk. As such, it can be conjectured that brain regions involved in encoding mean and risk in situations of pure risk (striatum, anterior cingulate cortex, insula, inferior frontal gyrus, …) are also involved in encoding these parameters when there is ambiguity. As the weights on mean and risk need not be the same, it is to be expected that the intensity of the brain activations differs depending on the situation.

This is supported by the findings of Hsu *et al*. (2005) and consistent with those of Huettel *et al*. (2006). The former found the striatum to be differentially activated under risk relative to ambiguity, whereas the amygdala and lateral orbitofrontal cortex showed the reverse. Furthermore, they provide both neuroimaging and brain-lesion data support for the idea that a common neural mechanism underlies both risk and

BOX 23.3

MODELS OF DECISION MAKING UNDER AMBIGUITY

A number of models exist that account for decisions under ambiguity. Perhaps the best known is the maxmin utility (MEU) model (Gilboa and Schmeidler, 1989), and its extension the α-maxmin (α-MEU) model (Ghirardato *et al*., 2004).

The former replaces the classic independence axiom with a weakened version (certainty-independence), thereby yielding the utility representation

$$MEU(f) = \min_{P \in C} \int_S u(f)dP,$$

where C is the set of probability measures on the set of possible states S. Under MEU, the decision maker considers only the worst-case scenario.

The α-maxmin model generalizes MEU by allowing the decision maker to consider a mixture of the worst and best-case scenarios, which provide anchor points for beliefs. That is,

$$\alpha MEU(f) = \alpha \min_{P \in C} \int_S u(f)dP + (1 - \alpha) \max_{P \in C} \int_S u(f)dP.$$

When $\alpha = 1$, beliefs are entirely determined by the worst-case scenario, and the αMEU utility model coincides with MEU. In the other extreme case ($\alpha = 0$), beliefs are solely based on the best-case scenario. For $\alpha > 0.5$, the decision maker is ambiguity averse; for $\alpha < 0.5$, she is ambiguity loving; at $\alpha = 0.5$, she is ambiguity neutral; her beliefs are the average of the worst-case and best-case scenarios.

ambiguity. First, the striatal activity was significantly correlated with the expected value of subjects' choices in both the risk and ambiguity conditions (Hsu *et al.*, 2005; see also Figure 23.7). Even stronger support is provided by the fact that patients with lesions to the lateral orbitofrontal cortex, as opposed to a comparison group of temporal-lobe lesion patients, were both risk *and* ambiguity neutral. Therefore, the difference between risk and ambiguity at the level of the reward system appears to be one of degree.

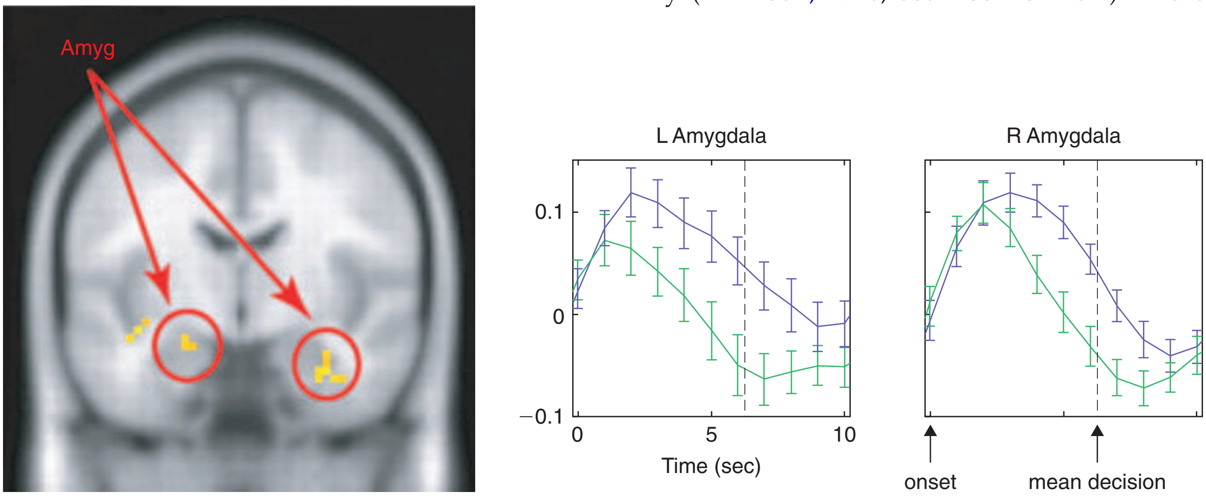
Closer inspection reveals that the mean-risk model is actually more general than the α-maxmin model; the latter imposes restrictions on the weights and on the mean-risk model. In fact, the mean-risk model could be derived as a hierarchical Bayesian updating model, whereby the decision maker is first agnostic about the true probability of winning (in the example above it is uniform between 1/3 and 1), but then uses the midpoint (the average between minimum and maximum probabilities) as a cue to where the true probability lies. Specifically, the midpoint is assumed to be drawn from a uniform distribution centered on the true probability and with a range equal to the difference between the maximum and minimum probabilities. The posterior mean and variance of the true probability are given by $\bar{p}$ and c, respectively.

The connection with updating and learning is important and illuminating in light of the findings of amygdala activation in Hsu *et al.* (2005). Significant activation of amygdala is rare or non-existent under pure risk. When contrasting choices involving ambiguity against those involving only pure risk, however, amygdala activation is strong. This could be consistent with activation of amygdala in, for example, novelty tasks (Rutishauser *et al.*, 2006), and suggests that the amygdala signals a need for learning of the unknown probabilities. Further research is needed to discover the true role of amygdala under ambiguity.

More generally, further experiments are needed to verify the extent to which valuation in the context of ambiguity is based on a computational model that trades off mean against risk. These experiments should extend previous studies pairing ambiguous with pure-risk (or zero-risk) gambles (Hsu *et al.*, 2005; Huettel *et al.*, 2006), and focus on choice between gambles with differing levels of ambiguity.

EXTENDING THE REWARD–RISK COMPUTATIONAL MODEL TO DECISIONS INVOLVING EQUITY AND EFFICIENCY

The separation between expected reward and variance we observe in neural valuation of risky gambles may also carry over to choices that require tradeoffs between efficiency and inequity. The latter is a typical problem in studies of distributive justice. A deep theoretical connection exists between decision making under uncertainty and the measurement of inequity (Atkinson, 1970; see also Box 23.4). The central



(a) (b)

FIGURE 23.7 Activation in amygdala is higher in free-choice trials involving an ambiguous option than those only involving pure risk. (a) Location of activation. (b) fMRI activation as a function of time since trial onset; dashed vertical line indicates mean time of choice; blue line is average activation in trials with ambiguity; green line is average activation in trials with pure risk; vertical line segments indicate 95% confidence intervals. Autocorrelation in time series is meaningless, as is scale of activation; only the difference in activation matters. From Hsu *et al.* (2005).

concern in both is the comparison of frequency distributions $f(y)$. In decisions under uncertainty, $f(y)$ denotes the probability of state y occurring, whereas in the case of inequity, $f(y)$ is the frequency distribution on income or other variables of interest for person y.

In experiments involving efficiency and inequity tradeoffs in a distributive justice task, Hsu *et al*. (2008) found a separation of regions encoding efficiency and inequity in much the same way as that for expected reward and variance. Specifically, efficiency is encoded in striatal regions including the caudate and putamen, whereas inequity is encoded in the insular cortex. This utilization of similar neural mechanisms for what on the surface are very different decisions provides a strong indication that these are very general mechanisms for the computation of decisions in a variety of different situations.

BOX 23.4

MEASUREMENT OF INEQUITY AND DECISION MAKING UNDER RISK

Central in the connection between measurement of inequity and decision making under risk is the idea of welfare. This is uncontroversial in the latter, and is traditionally meant to denote individual welfare, or utility. As pointed out in Dalton (1920), however, judgments of social welfare underlie the conception of any inequity measure. If we assume that the social welfare function is additively separable and symmetric in income, we arrive at the following:

$$W = \int_0^{\bar{y}} U(y) f(y) dy.$$

This form is immediately familiar in its resemblance to the standard expected utility representation. In fact, many of the concepts in decision making under risk, e.g., second-order stochastic dominance, mean-preserving spread, have formally identical counterparts in measurements of inequity, e.g., Lorenz dominance, principle of transfer, respectively (Atkinson, 1970). The assumptions of additive separability and symmetry, interestingly, can in addition be derived axiomatically via an appeal to decision making under risk (Harsanyi, 1978).

Acknowledgments

Financial support from the Swiss Finance Institute and the Swiss National Science Foundation (Finrisk NCCR) is gratefully acknowledged.

References

Agren, M. (2006). Prospect theory and higher moments. Working Paper, Uppsala University.

Atkinson, A.B. (1970). On the measurement of inequality. *J. Econ. Theory* 2, 244–263.

Bali, T.G., Demirtas, K.O., Levy, H. *et al*. (2008). Is there a relation between downside risk and expected stock returns? *J. Financial Quant. Anal.*, (forthcoming).

Bechara, A. and Damasio, A.R. (2005). The somatic marker hypothesis: a neural theory of economic decision. *Games Econ. Behav.* 52, 336–372.

Bechara, A., Damasio, H., Tranel, D. *et al*. (1997). Deciding advantageously before knowing the advantageous strategy. *Games Econ. Behav.* 275, 1293–1295.

Bechara, A., Damasio, H., and Damasio, A.R. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *J. Neurosci.* 19, 5473–5481.

Behrens, T.E., Woolrich, M.W., Walton, M.E. *et al*. (2007). Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221.

Bernheim, B. D. and A. Rangel (2008). Beyond revealed preference: toward choice-theoretic foundations for behavioral welfare economics. National Bureau of Economic Research Working Paper Series No. 13737.

Berns, G.S., Capra, M., Chappelow, J. *et al*. (2007). Nonlinear neurobiological probability weighting functions for aversive outcomes. *NeuroImage* 9, 2047–2057.

Blume, L. and Easley, D. (1992). Evolution and market behavior. *J. Econ. Theory* 107, 95–130.

Bruguier, A., Preuschoff, K., Quartz, S. *et al*. (2008). Investigating signal integration with canonical correlation analysis of fMRI brain activation data. *NeuroImage* 41, 35–44.

Camerer, C. and Weber, M. (1992). Recent developments in modeling preferences – uncertainty and ambiguity. *J. Risk Uncertainty* 5, 325–370.

Critchley, H., Mathias, C., and Dolan, R. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron* 29, 537.

Dalton, H. (1920). The measurement of the inequality of incomes. *Economic J.* 30, 348–361.

De Martino, B., Kumaran, D., Seymour, B. *et al*. (2006). Frames, biases, and rational decision-making in the human brain. *Science* 313, 684–687.

Diaconis, P. and Freedman, D. (1986). On the consistency of Bayes estimates. *The Annals of Statistics* 14(1), 1–26.

Dunn, B.D., Dalgleish, T., and Lawrence, A.D. (2006). The somatic marker hypothesis: a critical evaluation. *Neurosc. Biobehav. Rev.* 30, 239–271.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *Q. J. Economics* 75, 643–669.

Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902.

Ghirardato, P., Maccheroni, F., and Marinacci, M. (2004). Differentiating ambiguity and ambiguity attitude. *J. Econ. Theory* 118, 133–173.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *J. Math. Econ.* 18, 141–153.

Gomes, F.J. (2005). Portfolio choice and trading volume with loss-averse investors. *J. Business* 78, 675–706.

Hakansson, N.H. (1971). Multi-period mean-variance analysis: toward a general theory of portfolio choice. *J. Finance* 26, 857–884.

Harsanyi, J.C. (1978). Bayesian decision theory and utilitarian ethics. *Am. Econ. Rev.* 68, 223–228.

Hertwig, R., Barron, G., Weber, E.U. *et al*. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Sci.* 15, 534–539.

Hollerman, J.R. and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309.

Hsu, M. (2006). Three correlated essays on the neural foundations of economic decision making. PhD thesis, Faculty of Social Sciences, Pasadena, California Institute of Technology.

Hsu, M., Bhatt, M., Adolphs, R. *et al*. (2005). Neural systems responding to degrees of uncertainty in human decision making. *Science* 310, 1680–1683.

Hsu, M., Anen, C., and Quartz, S. (2008). The right and the good: distributive justice and neural encoding of equity and efficiency. *Science*, (in press)..

Huettel, S., Song, A., and McCarthy, G. (2005). Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *J. Neurosci.* 25, 3304–3311.

Huettel, S.A., Stowe, C.J., Gordon, E.M. *et al*. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron* 49, 765–775.

Kahneman, D. and Tversky, A. (1992). Advances in prospect theory: cumulative representation of uncertainty. *J. Risk Uncertainty* 5, 297–324.

Kahneman, D., Wakker, P.P., and Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *Q. J. Economics* 112, 375–405.

Lauwereyns, J., Takikawa, Y., Kawagoe, R. *et al*. (2002). Feature-based anticipation of cues that predict reward in monkey caudate nucleus. *Neuron* 33, 463–473.

Lucas, R. (1976). Econometric policy evaluation: a critique. *Carnegie-Rochester Conference Series on Public Policy* 1, 19–46.

Marinacci, M. (2002). Probabilistic sophistication and multiple priors. *Econometrica* 70, 755–764.

Markowitz, H. (1952). Portfolio selection. *J. Finance* 7, 77–91.

McClure, S.M., Berns, G.S., and Montague, P.R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38, 339–346.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In: P. Zaremba (ed.), *Frontiers in Econometrics*. New York, NY: Academic Press, pp. 105–142.

Montague, P.R., Dayan, P., and Schultz, W. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *J. Neurosci.* 16, 1936–1947.

O'Doherty, J.P., Dayan, P., Friston, K. *et al*. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron* 38, 329–337.

Paulus, M.P., Rogalsky, C., Simmons, A. *et al*. (2003). Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *NeuroImage* 19, 1439–1448.

Preuschoff, K. and Bossaerts, P. (2007). Adding prediction risk to the theory of reward learning. *Ann. NY Acad. Sci.* 1104, 135–146.

Preuschoff, K., Bossaerts, P., and Quartz, S.R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51, 381–390.

Preuschoff, K., Quartz, S.R., and Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *J. Neurosci.* 28, 2745–2752.

Rutishauser, U., Mamelak, A.N., and Schuman, E.A. (2006). Single-trial learning of novel stimuli by individual neurons of the human hippocampus–amygdala complex. *Neuron* 49, 805–813.

Segal, U. (1987). The Ellsberg Paradox and risk-aversion – an anticipated utility approach. *Intl Econ. Rev.* 28, 175–202.

Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642–1645.

Tobler, P.N., O'Doherty, J.P., Dolan, R.J. *et al*. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* 97, 1621–1632.

Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* 19, 181–189, (1 %R doi:10.1111/j.1460–9568.2004.03095.x).

Yu, A.J. and Dayan, P. (2003). Expected and unexpected uncertainty: ACh and NE in the neocortex. *Adv. Neural Inf. Proc. Syst.* 15, 157–164.

CHAPTER

24

Multiple Forms of Value Learning and the Function of Dopamine

Bernard W. Balleine, Nathaniel D. Daw, and John P. O'Doherty

OUTLINE

INTRODUCTION

Among the key findings in the behavioral psychology and systems neuroscience of decision making is that the same behavior – for instance, a rat's lever-press – can arise from multiple influences that are both neurally and psychologically dissociable. In this chapter, we review recent theory and research about these decision mechanisms and their neural bases, focusing particularly on the idea that they embody distinct evaluative or motivational processes – that is, different sorts of "value".

Although there is an extensive literature linking the control of executive functions to the prefrontal cortex (Goldman-Rakic, 1995; Fuster, 2000), more recent studies suggest that these functions depend on reward-related circuitry linking cortex with the striatum (Chang *et al.*, 2002; Lauwereyns *et al.*, 2002; Tanaka *et al.*, 2006). Evidence from a range of species suggests that discrete cortico-striatal networks control two functionally distinct decision processes. The first involves actions that are more flexible or *goal-directed*, sensitive to reward-related feedback, and involve regions of association cortex – particularly medial prefrontal, medial orbitomedial, premotor, and anterior cingulate cortices together with their efferent targets in caudate/dorsomedial striatum (Haruno and Kawato, 2006; Levy and Dubois, 2006). The second involves actions that are relatively automatic or *habitual*, and depend on sensorimotor cortices and dorsolateral striatum/putamen

(Jog *et al.*, 1999; Poldrack *et al.*, 2001). These two types of processes have been argued to employ different learning rules (Dickinson, 1994), different forms of plasticity (Partridge *et al.*, 2000; Smith *et al.*, 2001), and different computational principles (Dayan and Balleine, 2002; Daw *et al.*, 2005). Furthermore, degeneration of these two cortico-striatal circuits has been argued to result in two distinct forms of pathology, such as Huntington's disease, obsessive compulsive disorder and Tourette's syndrome on the one hand (Robinson *et al.*, 1995; Bloch *et al.*, 2005; Hodges *et al.*, 2006) and Parkinson's disease and multiple system atrophy on the other (Antonini *et al.*, 2001; Seppi *et al.*, 2006).

Interestingly, these distinct mechanisms of action control appear to be related to distinct motivational or evaluative processes. Goal-directed actions are so named because they are sensitive to feedback about what we shall call *goal values* – that is, the rewarding values of the particular outcomes (for instance, the specific sort of food delivered for a lever-press). Further demonstrating their goal sensitivity, such actions are also sensitive to predictive cues that signal the likelihood of those outcomes independent of any actions – referred to here as *Pavlovian values*. In contrast, habitual actions have long been argued to arise from a generalized propensity for emitting a particular response, which is acquired from and grounded in a history of reinforcement but divorced from any representation of the specific reinforcing goal. We refer to these generalized action propensities as *habit values*.

This chapter will review the behavioral evidence for these distinct value functions, computational models that relate these functions to distinct aspects of adaptive behavior, and the neural systems that mediate these processes. Of particular interest will be the role of the midbrain dopamine system, and the extent to which it is involved (separately or integratively) in these valuation processes.

REWARD, PREDICTION, AND REINFORCEMENT

A classical economic notion of choice assumes that people evaluate options according to their expected utilities. Human and animal actions, however, are not nearly so unitary.

Psychologists (see Box 24.1) have long attempted to distinguish *bona fide* decisions from other behaviors, such as reflexes, that might only appear to be choice-like. In one view, for an action to qualify as truly volitional or "goal-directed" it should depend on two factors, which echo the probability and utility of an outcome from the standard formula for expected utility. The first is knowledge of the *contingency* between the action and some outcome; the second is the *valuation* of that outcome as a desirable goal.

As detailed below, these two criteria have been operationalized into behavioral tests. These tests reveal a clear dissociation between behaviors that are demonstrably goal-directed in this sense, and a second class of "habitual" behaviors, which pass neither test. As sketched more mathematically under "Reinforcement learning", below, these behaviors can be viewed as arising from different processes for *evaluating* an action – either through something like an explicit computation of expected utility, or through a shortcut that simply assumes that previously reinforced behaviors will be valuable.

We begin by describing evidence for yet a third class of behaviors, and underlying valuations, which also fail to satisfy the criteria for goal-directed actions. These are the conditioned reflexes studied by Pavlov (1927).

Reflex Versus Action

Although most aspects of our behavioral repertoire can be described with respect to attaining some goal or other, many of these activities are actually simple reflexes that are elicited without deliberation by environmental cues. As discussed above, a critical distinction between reflexive responses and goal-directed actions is that the latter are controlled by knowledge of their relationship to their consequences whereas the former are not. Classic examples of reflexive responses that have a misleading veneer of choice-like goal-directedness about them can be found in simple conditioning situations, such as that made popular by Pavlov (1927). He studied salivary responses to food delivery in dogs, and the conditioning of those reflexes produced by pairing a neutral stimulus (such as a tone) with food.

The key feature of this sort of task is that food has a contingent ("*Pavlovian*") relationship with the stimulus, but its delivery is not contingent on any action the animal takes. The fact that the stimulus often acquires the ability to elicit anticipatory salivation in this situation is usually thought to reflect the transfer of control over the salivary reflex from the food to the stimulus, based on the Pavlovian association between stimulus and food (see Figure 24.3a, later in this chapter).

From a decision-making perspective, however, it is possible to argue that in fact in this situation dogs control their salivation and *decide* to produce this

BOX 24.1

BASIC LEARNING PROCESSES AND SOME TERMINOLOGY

Psychologists have long studied fundamental learning processes. To date, these have been isolated as being those related to (1) *stimulus habituation*, i.e., the reduction in the response elicited by a stimulus that is induced by its repeated presentation; (2) *Pavlovian* or *classical conditioning*, in which the delivery of a biologically potent event (or US; i.e., food, water, a predator and so on) is predicted by, or made conditional upon, a relatively neutral stimulus (or CS) and, as a consequence, the reflexive unconditioned response (UR) induced by exposure to the US comes to be elicited by the CS (for example, contact with food can elicit salivation, and so can a bell that has been reliably paired with food); and (3) *instrumental conditioning*, in which access to a motivationally valuable commodity (for example, food when hungry; safety when apprehensive, etc.) is made conditional on the performance of the animal's own actions, such as pressing a lever or pulling a chain.

Various aspects of these types of learning should be recognized. First, it is possible to talk descriptively about learning in terms of the acquisition of a response. It is also possible to talk in terms of the mechanism that mediates that change in response. These have resulted in both behavioral and cognitive definitions and theories of learning. In a behavioral definition, all learning involves the association of stimuli (S) and behavioral responses (R) of one kind or another (i.e., all learning is S–R). On cognitive views, learning involves the formation of a novel cognitive structure or association – for example, between stimuli (CS–US) or between responses and outcomes (R–O) that are then manifest in performance (of the CR in the case of CS–US, and of R in the case of R–O associations). As discussed in this chapter, much evidence supports the cognitive perspective, but not all. Some training conditions appear to be particularly apt for producing S–R learning.

Finally, a number of general terms are used here that refer to commonly-used procedures:

Contingency: The net probability of one event given another (for example, of a US given a CS, or of an outcome given a response) – i.e., the difference between the probability of event 2 given event 1 ($P(\text{E2}\,|\,\text{E1})$) and of event 2 given no event 1 ($P(\text{E2}\,|\,\text{noE1})$). There is a positive contingency when $P(\text{E2}\,|\,\text{E1}) > P(\text{E2}\,|\,\text{noE1})$ and a negative contingency when $P(\text{E2}\,|\,\text{E1}) < P(\text{E2}\,|\,\text{noE1})$.

Outcome revaluation: A procedure that serves to change the motivational or reward value of an event. For example, if rats are trained to respond for sugar, the sugar could be revalued by increasing its value by, say, an increase in food deprivation, or by decreasing its value by a decrease in food deprivation. Sugar can also be devalued using taste-aversion procedures, such as pairing the sugar with the induction of illness (see Figure 24.1).

Extinction: This refers to the reduction in the performance of a CR or of an instrumental response when the CS or the instrumental action is no longer followed by the US or outcome with which it was previously paired.

response, perhaps to facilitate digestion or to improve the taste of food. To assess this explanation of conditioned salivation, Sheffield (1965) arranged a standard Pavlovian conditioning experiment in which he paired a tone with food delivery, but with a twist: if the dog salivated during the tone, then the food was not delivered on that trial. This arrangement maintains a Pavlovian relationship between the tone and food, but abolishes any positive relationship between salivation and food. Sheffield reasoned that if the salivation was an action controlled by its relationship to food, then arranging that salivation omitted food delivery should ensure that the dogs stop salivating – indeed, having never had the opportunity to learn that salivating improved the reward value of the food by enhancing its flavor or improving its ingestion, they should never acquire salivation to the tone at all. Sheffield found that it was clearly the Pavlovian tone–food relationship that controlled salivary performance; during the course of over 800 tone–food pairings, the dogs acquired and maintained salivation to the tone even though this resulted in them losing most of the food they could have obtained by withholding their salivary response.

Although salivation might in principle be the exception, numerous other studies over the past 40 years have established, in a range of species including humans (Pithers, 1985) and other animals (Williams and Williams, 1969; Holland, 1979), that a large variety of conditioned responses are maintained by the relationship between the Pavlovian predictive cue and the outcome (food, shock, etc.) with which it is associated, rather than by the relationship between the response and the outcome.

Pavlovian Values Versus Goal Values

Though Pavlovian responses do not satisfy the criteria for goal-directed action, the same is not true of behaviors acquired in other conditioning preparations, notably instrumental conditioning (Thorndike, 1911). Here, animals readily learn to produce new and often quite arbitrary responses (such as lever-presses) to gain access to food, which, unlike in Pavlovian conditioning, is delivered contingent on the action. In contrast to the salivation response studied by Sheffield, putting these responses on an omission contingency, in which responding leads to the omission of an otherwise freely delivered food, rapidly reduces their performance (Davis and Bitterman, 1971; Dickinson *et al.*, 1998). Furthermore, numerous studies have demonstrated the exquisite sensitivity of the performance of instrumental actions to changes in the net probability of outcome delivery given the action (i.e. the difference between the probability of food given a response and the probability of that food given no response). These changes can be highly selective; for instance, in situations in which different actions lead to different outcomes, degrading a particular action-outcome contingency by delivering its associated outcome noncontingently often has no effect on the performance of other actions (Colwill and Rescorla, 1986; Dickinson and Mulatero, 1989; Balleine and Dickinson, 1998). Like the omission test, this contingency degradation test exercises the first of the criteria for goal-directed action discussed above – that it be based on knowledge of the action-outcome contingency (Figure 24.3a).

Instrumental responses are sensitive not only to the action-outcome contingency, but also to the second criterion for a goal-directed action: the value of the goal. Again, experiments in rats have provided some of the clearest demonstrations of this effect. For example, in a number of studies, illustrated in Figure 24.1, hungry rats have been trained to carry out two actions (presses of different levers, or a lever-press and a chain-pull), with one response earning, say, food pellets, and the other a sucrose solution. After this training, the

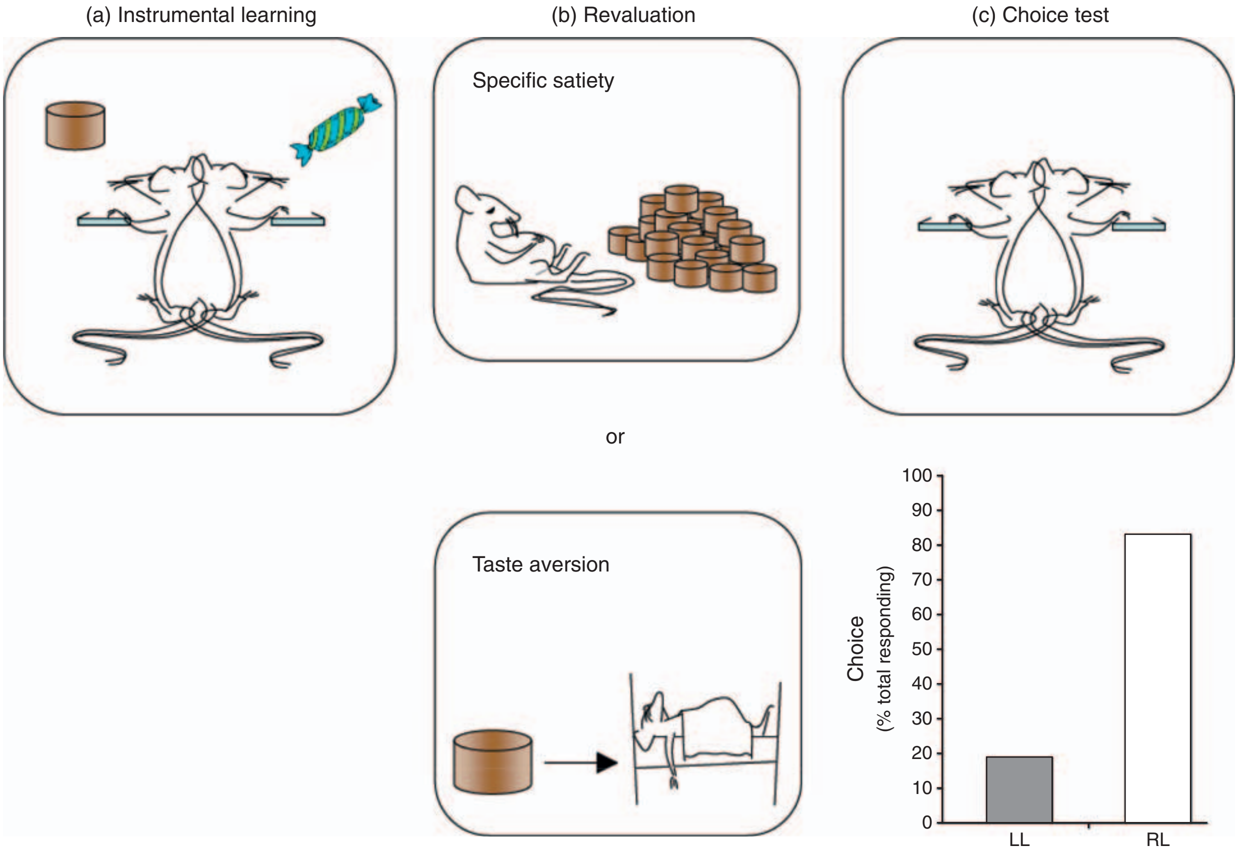


FIGURE 24.1 Assessing the influence of goal values on choice, using outcome devaluation. (a) Rats are first trained to perform two actions, each earning one of two foods. (b) One of the two foods is then revalued; here this is illustrated either by taste-aversion learning (lower panel) or by a specific satiety treatment (upper panel). (c) The influence of these treatments on choice is then assessed in a choice test conducted in extinction (i.e. the absence of either food outcome). (d) Typically, devaluation biases choice away from actions (LL: left lever) that, in training, earned the devalued outcome, and toward those (RL: right lever) that earned outcomes that were not devalued. Data redrawn from Corbit and Balleine (2003), with permission.

desirability of one of the two outcomes is reduced, either by specific satiety (in which the rats are allowed to eat a large quantity of one outcome) or, in other studies, by taste-aversion learning (in which consumption of one outcome is followed by drug-induced illness; see Figure 24.1). Both treatments reduce the consumption of a specific food relative to other foods. After the devaluation treatment, the rats are given a choice test between the two actions *in extinction* (i.e., in the absence of the delivery of either food outcome). If – as expected for a goal-directed behavior – performance of the actions is maintained by the values of their respective outcomes (their *goal values*), then devaluation of one outcome should reduce subsequent performance of its associated action in the extinction test, relative to the other action. This is exactly the result that is commonly reported; devaluing one of the food rewards selectively and profoundly reduces the performance of the action that in training delivered that outcome, compared to the action trained with the non-devalued outcome (Dickinson and Balleine, 1994; Balleine, 2001).

Instrumentally trained actions, unlike Pavlovian conditioned reflexes, therefore satisfy both criteria for goal-directed decisions. However, this should not be taken to indicate that Pavlovian learning has no effect on decision making. To the contrary, instrumental behaviors are also potently affected by Pavlovian cue–outcome associations, as demonstrated in *Pavlovian-instrumental transfer* studies. These show that choice between instrumental actions, which (as reviewed above) is sensitive to the goal value and contingency, can also be substantially modified by the presentation of reward-related Pavlovian cues.

For example, in a recent study, illustrated in Figure 24.2, Corbit and Balleine (2005) exposed hungry rats to Pavlovian training in which two distinct auditory cues, a tone and white noise, were used to predict the delivery of two different but equally valued foods – grain food pellets and a sucrose solution. After this phase, the rats were trained to push two levers, one paired with the grain pellets and the other with the sucrose solution.

In short, the first phase trains rats on a Pavlovian contingency (cues with outcomes, and no lever-pressing) and the second phase on an instrumental contingency (lever-pressing for outcomes, without the tone or noise). The third phase tests the combined effect of both contingencies, by giving the rats a choice test in which the rats are allowed freely to choose between the two levers during presentations of the tone and noise stimuli. During this test, neither the grain pellet nor the sucrose outcomes were delivered, so any performance could only be attributable to the previous learning. Although the rats' performance on the two levers was the same in the absence of the auditory stimuli, presentation of the tone was found to increase the rats' choice of the lever that, in training, had delivered the same outcome as the tone, whereas presentation of the noise was found to increase the rats' choice of the other lever (see Figure 24.2). Numerous demonstrations of this effect have been reported, confirming that choice between actions is determined not just by goal values but also by the specific predictions – what we refer to here as the *Pavlovian values* – provided by Pavlovian, reward-related cues (Colwill and Rescorla, 1988; Colwill and Motzkin, 1994; Corbit *et al.*, 2001; Corbit and Balleine, 2003).

Note that, from a pure utility-based decision making perspective, this effect is somewhat puzzling, because in training the Pavlovian cues predict outcome delivery not contingent on any action, and it is therefore unclear why their predictions should be relevant to the valuation of actions. One clue comes from an experiment by Delamater (1995), who trained to a cue with Pavlovian predictive associations (e.g. noise predicting sucrose) but then selectively degraded its contingent predictive relationship by presenting the sucrose both during the noise and during a period when the noise was not presented. The degraded stimulus no longer showed Pavlovian-instrumental transfer; that is, its presentation failed to produce a selective increase in pressing a lever that had previously been paired with sucrose. This result is important in demonstrating that it is not just the pairing of a stimulus with an outcome that mediates its effects on choice; it is the *information* that the stimulus provides as a net predictor of a specific rewarding event. That the effect of a Pavlovian cue on instrumental choices depends on its predictive validity suggests that animals interpret it as providing information regarding the availability or likelihood of achieving a specific goal.

We have reviewed evidence that instrumental choice is biased both by goal values, as demonstrated in outcome devaluation studies, and also by Pavlovian cue-outcome values in Pavlovian-instrumental transfer. One question is whether these two effects really reflect distinct evaluative processes, or whether they instead are somehow actually mediated by the same mechanism. This might be true, for instance, if sensory cues related to the actions themselves play a role analogous to the lights and tones in the Pavlovian case. Although views of this sort have been formalized in a number of versions (Rescorla and Solomon, 1967; Trapold and Overmier, 1972), there is now considerable evidence confirming that goal values and Pavlovian values are mediated by distinct processes (Balleine and Ostlund, 2007). For instance, the strength of the two effects – the effect of Pavlovian cues in

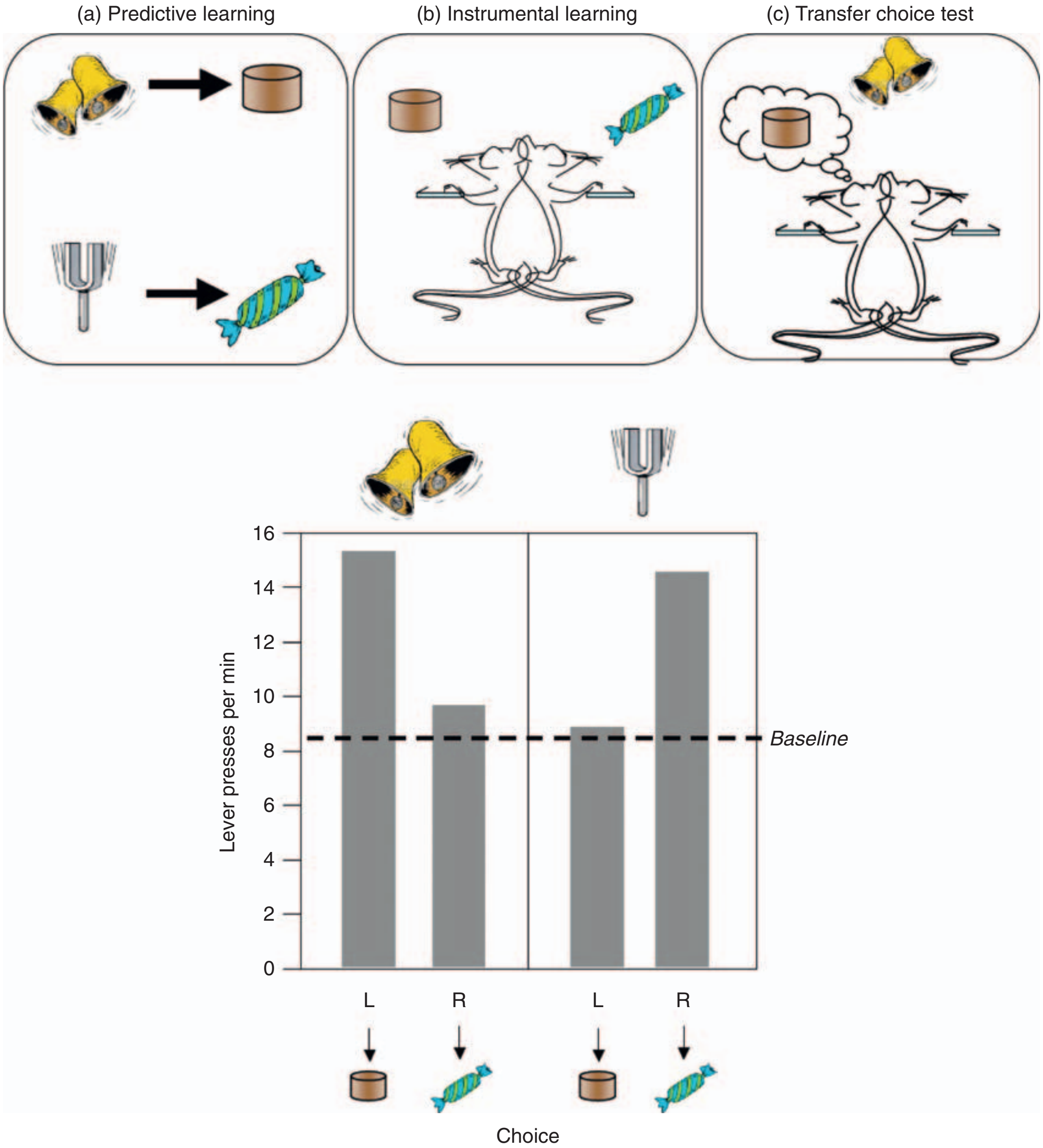


FIGURE 24.2 Assessing the influence of Pavlovian values on choice, using Pavlovian–instrumental transfer. (a) Rats are first trained to predict two food outcomes based on distinct auditory cues and then (b) to perform two actions each earning one of the two foods. (c) The influence of the predictive cues on choice is assessed in a transfer choice test conducted in extinction (i.e. the absence of either food outcome). (d) Typically, a stimulus biases choice towards actions that earn the outcome predicted by the stimulus. Data redrawn from Corbit and Balleine (2003), with permission.

transfer experiments and the effect of outcome devaluation – can be independently, differentially modulated by several behavioral manipulations, and one can be observed in circumstances when the other is not present (Corbit and Balleine, 2003; Holland (2004).

Goal Values Versus Habit Values

Importantly, only some instrumentally trained actions, such as newly acquired lever-presses, demonstrate the two defining features of goal-directed actions in the manner detailed above. Other instrumental behaviors, such as the same responses following overtraining, can become more automatic, involuntary, or impulsive, and fail to satisfy these tests – i.e., they become *habits* (Dickinson, 1994). There has been considerable interest in habits as a putative counterpart in normal function to the loss of behavioral control in various neurodegenerative conditions and in drug addiction (Dickinson *et al.*, 2002; Robbins and

Everitt, 2002; Miles *et al.*, 2003; Cardinal and Everitt, 2004). These ideas are rooted in classic theories of stimulus–response (S–R) reinforcement learning (Hull, 1943). According to these, rewarding events reinforce or create direct, causal associations between contiguous sensations and responses, allowing the stimulus directly to elicit the response in a manner that is no longer dependent on the response-outcome contingency or the outcome's value (Holman, 1975; Adams and Dickinson, 1981; Dickinson *et al.*, 1983; Dickinson *et al.*, 1995) (Figure 24.3a).

Although it is straightforward to apply these ideas to drug addiction, only relatively recently was direct evidence found to suggest that these mechanisms also apply to activities associated with natural rewards like food. For example, to test contingency sensitivity, Dickinson *et al.* (1998) trained hungry rats to press two levers for food pellets before delivering a sugar solution freely and periodically. Responding on one lever had no effect on sugar delivery, but responding on the other delayed it; in other words, to maximize their access to both food pellets and sugar, the rats had to withhold responding on one lever but not the other. Consistent with the results discussed in the previous section, this proved to be a relatively easy task for undertrained animals. However, animals who had been overtrained on the initial lever-pressing task did not adjust readily to this omission contingency, and kept responding similarly on both levers even though this lost them significant access to the sugar (Dickinson *et al.*, 1998). This same effect has been replicated in mice (Frankland *et al.*, 2004).

After overtraining, instrumental actions can also fail to satisfy the second criterion for goal-directed behavior: they can be insensitive to changes in goal value. Holman (1975) showed that overtrained lever-presses in thirsty rats reinforced by access to a saccharin solution persisted even after the saccharin had been devalued by pairing its consumption with illness. (The test was performed in extinction – that is, without further delivery of the saccharin.) It is important to recognize how maladaptive the lever-pressing was in Holman's rats. Although the pairing with illness resulted in the rats no longer consuming or even contacting the previously palatable (but now poisonous) saccharin, they continued to work on the lever at a rate comparable to that of rats for which the saccharin was not devalued. This effect was replicated several times over the next decade (Adams, 1981; Adams and Dickinson, 1981; Dickinson *et al.*, 1983). These findings provided direct evidence that, in addition to control by a goal-directed process, the performance of instrumental actions can also become habitual. The later experiments also show that either process can be engaged depending not just on the degree of training but also on the relationship between instrumental performance and reward delivery. In particular, when rewards are delivered on a timed schedule so that changes in the rate of lever-pressing have little effect on the rate of reward, actions tend to become habitual. When rewards are delivered so that the experienced rate of reward is proportional to the rate of lever-pressing, however, actions tend to remain goal-directed.

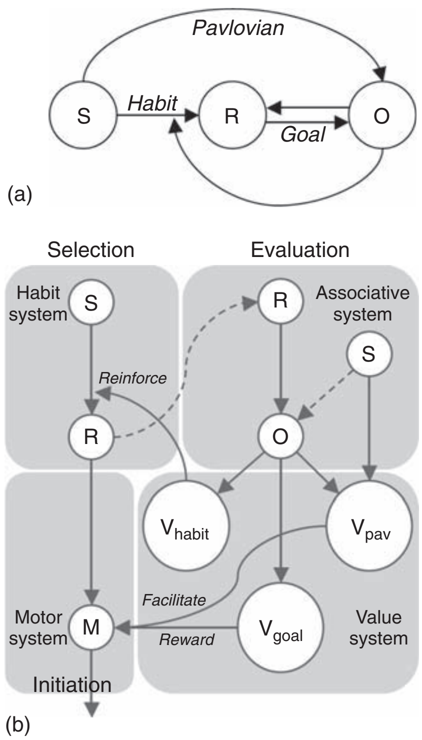


FIGURE 24.3 The associative structure of instrumental action. (a) A schematic overview summarizing goal, habit, and Pavlovian values and their influence on action through associations between stimuli, responses, and outcomes. (b) Process view of instrumental action: the associative basis of action selection, evaluation and initiation. In this model, action selection is modulated by outcome value and inputs from both selection and evaluative processes are required for actions to be performed. Performance can be facilitated by Pavlovian values and, after a period of invariant training, the increased habit value can render it sufficiently strong for selection to result immediately on action initiation, without feedback from the goal value.

To review, the experiments above suggest that there exists a second class of "habitual" instrumental actions that are distinguished from goal-directed actions by failing to meet the criteria of contingency or outcome sensitivity. As mentioned above, these habitual responses are classically envisioned to arise from stimulus–response associations, which lack any representation of the rewarding outcome and are instead valued in the sense of having been "stamped in" by a history of reinforcement (Hull, 1943). We refer to this propensity to emit a habitual response as a *habit value*. Finally, and interestingly, animals not only acquire new actions in situations where they are reinforced by primary rewards such as food, but also when they are reinforced by stimuli associated with those primary rewards. This latter effect, called *conditioned reinforcement*, has also been argued to be mediated by a process of S–R association, in this case reinforced by the stimulus signaling reward rather than the reward itself. In confirmation of this suggestion, recent studies have found evidence that the actions acquired through this process are insensitive to outcome devaluation (Parkinson *et al*., 2005).

Expectancy, Reinforcement, and Reward

As described in this section, behavioral research provides considerable evidence for at least three distinct forms of evaluative process, which bias decision making based on expectancy, reinforcement, and reward (Figure 24.3b). Expectancy induced by stimuli that predict reward – Pavlovian values – can exert selective effects on action selection; these stimuli selectively elevate the performance of actions associated with the same outcome as that predicted by the stimulus, relative to other actions. Only cues that are valid predictors of an outcome associated with an action affect its performance in choice tests, suggesting that this effect is based on the predictive information that the stimuli provide about the likelihood of specific outcomes.

This effect of Pavlovian values differs from the biasing of action selection induced by reinforcement, whether it is an effect of primary or of conditioned reinforcement. In this case, the bias toward the performance of an action that has proven to be successful in the past is based on an association between the action and the stimulus or state in which the action was rewarded. Hence, this is considered a reinforcement function of reward delivery; it is not the anticipation of the reward but the strengthening of the stimulus–response association that produces the change in performance of the response and establishes the strength of the habit value.

Finally, the reinforcing function of reward delivery is to be contrasted with its function as a valued goal of an action. The reward value of goals clearly biases the animals' tendency to select and initiate a particular course of action in a way that does not depend upon the state or stimulus with which that goal is associated, or on the function of the goal as a reinforcer of a specific stimulus–response connection. Rather, reward value appears to modulate the tendency to initiate an action that is selected based on the encoding of the action-outcome contingency. These distinct functions are best illustrated in the differential effects of reward devaluation and contingency degradation on undertrained and relatively more overtrained actions.

Generally, therefore, as illustrated in Figure 24.3b, behavioral evidence points to three dissociable "value" functions that guide the choice and decision-making process. But how deep does this analysis run? In the following sections, we will assess this question from the perspective of both contemporary computational models and their neural implementation within the functional systems that current evidence suggests mediate the neural bases of decision making.

REINFORCEMENT LEARNING

In contrast to psychologists studying natural agents, computer scientists have also long studied issues of prediction and action in the context of controlling artificial agents such as robots. In fact, there have been important relationships forged between the approaches taken to the study of adaptive behavior in natural and artificial systems. Reinforcement learning (Sutton and Barto, 1998), for instance, is the study of how to learn, by trial and error, to make efficient decisions. Apart from its practical engineering uses, reinforcement learning has provided an apt set of models for both conditioning behavior and the underlying neurophysiology.

Here, we define formally the problem of reinforcement learning, and review some basic approaches to solving it (for more details, see also Chapter 21 of this volume). In the process, we will note striking analogies between these computational tools and several of the psychological constructs outlined above, notably those of Pavlovian, habitual, and goal values.

The Markov Decision Process

Most work in reinforcement learning targets a class of model tasks, the Markov decision process (MDP), which is simplified enough to admit of formal analysis but still embodies many features of non-trivial real-world decisions.

In an MDP, the counterpart to a stimulus in conditioning is a *state*. At each timestep t, the model world takes on a discrete state s_t. Also, the agent chooses some action a_t. Together, these probabilistically determine the next state, s_{t+1}. Formally, the structure of a particular MDP is defined by the *transition function*

$$T(s,a,s') = P(s_{t+1} = s' \mid s_t = s, a_t = a) \quad (24.1)$$

which specifies the probability distribution over the new state (s') that follows any state (s), action (a) pair.

In selecting actions, the goal of the agent is to obtain rewards. The variable r_t measures (as a real number) the reward value that the subject receives on trial t. We assume that this is a function of the state: $r_t = R(s_t)$. Together, the *reward function* R and the transition function T define an MDP.

For instance, a typical conditioning experiment might be modeled with states corresponding to the different controlled stimuli and outcomes, actions such as lever-presses, a transition function detailing how actions bring about outcomes, and, finally, a reward function mapping some of these states (e.g. those corresponding to consuming various foods like pellets and sucrose) to their rewarding value. Note that this formalism respects the distinction between the predicted sensory and the motivational aspects of an outcome (here, represented by the outcome state itself and its reward value). The following sections describe how different sorts of value can be built up from these elements. As with their psychological counterparts, goal and habit values, these can be differentially sensitive to manipulations of reward value, for example, by illness or satiety treatments. In the present setting, these revaluation manipulations can be viewed as changing the reward function (Daw *et al.*, 2005; Niv *et al.*, 2006).

Action Selection in MDPs

The central problem in reinforcement learning is being dropped into an unknown MDP – like a rat into a Skinnerbox – and learning, by trial and error, how best to gather rewards. A major difficulty in this setting is that each action influences the next state and reward, but this state also influences the next one and its reward, and so on. We must therefore evaluate actions in terms of their long-term consequences.

To this end, we can define the *state-action value function*, Q, as the cumulative reward that follows a particular action in a particular state:

$$Q(s,a) = E[r_t + r_{t+1} + r_{t+2} + r_{t+3} + \ldots \mid s_t = s, a_t = a] \quad (24.2)$$

Here, the expectation $E[\cdot]$ averages over randomness in the state transitions, and we are being temporarily ambiguous about what actions the agent takes at times $t + 1$ and thereafter.

The motivation for this definition is that since $Q(s,a)$ measures the long-term reward following an action, it ameliorates the principal difficulty of decision making in an MDP. If we knew it, then at any state optimal action selection would be as simple as choosing whichever action had maximal value. The strategy is to guide choice by predicting future rewards.

We can rewrite the function in a more useful form by noting that the bulk of the sum (all the terms starting with r_{t+1}) is just the value of the successor state, $Q(s_{t+1}, a_{t+1})$. This motivates the following *recursive* definition of the long-term value Q in terms of itself:

$$Q(s,a) = R(s) + \sum_{s'} T(s,a,s') \max_{a'}[Q(s',a')] \quad (24.3)$$

Here, the first reward r_t is determined by the MDP reward function R, and the average over the successor state is written explicitly, according to the probabilities given by the transition function T. Also, this equation resolves the ambiguity about what action the agent takes in the successor state s' by assuming he takes the best action there, which has value $\max_{a'}[Q(s', a')]$, that is the value of the best action a' in that state.

Formally, the definition of long-term value in equation (24.3) specifies what it means to choose optimally in an MDP. More practically, it serves as the basis for a number of different approaches to reinforcement learning. These simply correspond to different methods of using experience to solve equation (24.2) for the action value function $Q(s,a)$.

Returning finally to the psychological context, we may ask to which of our notions of value this long-term value corresponds. In fact, we will suggest, it corresponds to *both* habit value and goal value. In this view, the difference between the two is not what information they carry, but rather in how this information is inferred from experience.

In particular, estimates of Q may be derived using different reinforcement learning procedures that either do or do not make use of an intermediate representation of the action–outcome transition function and the identity of the goal outcome. These two sorts of value suggest computational counterparts to goal and habit values, respectively, in that the first representation satisfies both of criteria for goal-directedness from the previous section, whereas the latter does not (Daw *et al.*, 2005).

Model-based Reinforcement Learning and Goal Values

A straightforward, though indirect, approach to predicting action values in an initially unknown task is simply to attempt to figure out what the task is: formally, to estimate the reward function R and transition function T that together define the MDP. These functions form a complete description, or *model*, of the MDP, and methods that work by recovering them are accordingly known as *model-based* reinforcement learning. Note that these functions are easy to learn from experience; each state transition observed or reward received is a direct sample from one of these functions, and one need only average this raw experience to estimate the underlying functions.

Given R and T, it is possible, though computationally laborious, to estimate the action value function $Q(s, a)$ simply by evaluating equation (24.3). The standard procedure for doing this is called *value iteration*. It requires plugging in these functions, then iteratively expanding the future value term on the right-hand side of equation (24.3), adding up expected rewards over different anticipated trajectories of potential future states and actions. The process is rather like how a novice evaluates chess moves, by exhaustively considering series of future states (board positions) and actions (moves and countermoves).

Values derived this way are a computational counterpart to goal values (Daw *et al.*, 2005). This is because the transition function T embodies knowledge, analogous to response–outcome associations, about how particular actions in particular states lead to other states, including those representing particular outcomes like various foods. Similarly, the reward function R represents the rewarding value of those outcomes. Decisions derived from these functions will therefore be sensitive to outcome contingency and value, the two criteria for goal-directed actions. For instance, if the reward value of these outcomes changes (as represented by the reward function R), then Q values and choices derived from this knowledge will immediately reflect the change.

Model-free Reinforcement Learning and Habit Values

Instead of estimating future value indirectly, via the transition and reward functions, an alternative approach is to estimate it directly, by averaging samples of the right-hand side. One version of this idea, called *Q-learning* (Watkins, 1989), works as follows. Suppose we have some initial estimate of the value function, which we denote $\hat{Q}(s, a)$ to distinguish it from the true function. On timestep t, starting in state s_t, we receive reward r_t and take action a_t. This leads us to a new state s_{t+1}, drawn from the transition distribution $T(s_t, a_t)$. This small snatch of experience gives us a new estimate of the future value $\hat{Q}(s_t, a_t)$, to wit: $r_t + \max_{a'}[\hat{Q}(s_{t+1}, a')]$, and we may update our previous estimate $\hat{Q}(s_t, a_t)$ by averaging this in. In the next step, we may go on to update our estimate $\hat{Q}(s_{t+1}, a_{t+1})$ using the subsequent reward and state transition, and so on.

A standard approach is to mix the new sample with the old prediction in proportions determined by a "learning rate" parameter η. The resulting update rule for $\hat{Q}$ can then be written as:

$$\hat{Q}(s_t, a_t) \leftarrow \hat{Q}(s_t, a_t) + \eta\delta_t \qquad (24.4)$$

where the "prediction error" δ_t measures the difference between the old prediction and the newly observed sample:

$$\delta_t = r_t + \gamma \max_{a'}[\hat{Q}(s_{t+1}, a')] - \hat{Q}(s_t, a_t) \qquad (24.5)$$

Algorithms of this sort are known as *temporal-difference* algorithms, because the last two terms of the prediction error are the difference between estimates $\hat{Q}$ made at successive timesteps. They are also known as *model-free* because, in contrast to the approach discussed in the previous section, they do not rely on any representation of the MDP.

For this reason, like habit values, values predicted this way will be insensitive to changes in the reward value (Dayan and Balleine, 2002; Daw *et al.*, 2005). Since the learned function $\hat{Q}(s, a)$ is not grounded in any information about the *identities* of the future outcomes (such as food) contributing to the predicted value, it is analogous to the formation of a stimulus–response association. Also, like habit values, values learned this way are grounded in a history of reinforcement rather than a prediction of a particular future reward.

The Actor/critic and Pavlovian Values

One variation on temporal-difference methods is particularly relevant to the study of conditioning because it decomposes value learning into learning about states ("Pavlovian" learning) and, separately, learning about actions ("instrumental" learning). These methods are known as actor/critic algorithms (Barto, 1995).

A reinforcement learning counterpart to Pavlovian value is a function measuring future rewards expected from a state, while ignoring actions. (That is, strictly speaking, averaging out the agent's action choices as though they were just another source of randomness in state–state transitions. The values derived are denoted as "V" values rather than "Q" values).

The state value function has its own recursive definition, analogous to equation (24.3):

$$V(s) = R(s) + \sum_{a}\sum_{s'} \pi(s,a)T(s,a,s')V(s') \quad (24.6)$$

(where $\pi(s, a)$ is the probability that action a will be chosen in state s). It also has its own temporal-difference learning rule with error signal

$$\delta_t^V = r_t + \hat{V}(s_{t+1}) - \hat{V}(s_t) \quad (24.7)$$

Given a Pavlovian module that learns to predict $\hat{V}(s)$, called a *critic*, we can define a separate *actor* module that uses this information to learn how to select actions. One version of this – called "advantage learning" (Baird, 1994, Dayan and Balleine, 2002) – defines the *advantage* A of action a in state s as

$$A(s,a) = Q(s,a) - V(s) \quad (24.8)$$

that is, the difference between the future value of taking action a and the future value averaged over actions according to their usual frequencies π. As it turns out, this function can be estimated simply by averaging TD errors from equation (24.6) (which are samples of the advantages). Here, like Q, $\hat{A}(s,a)$ serves as a habit value implicitly defining an action selection policy (take whichever action has the highest advantage at a state); the Pavlovian predictions from the critic help to learn this policy by defining the error signal δ_t^V that trains it.

Actor/critic algorithms like this one work by learning Pavlovian predictions $\hat{V}(s)$ and then learning to choose those actions that tend to lead to states with a high Pavlovian value. The psychological counterpart is conditioned reinforcement, in which responses can be reinforced not just by primary rewards but also by Pavlovian cues predicting reward. As discussed above, this is only one of the ways in which Pavlovian and instrumental conditioning interact. Other, more selective, effects of Pavlovian values on choice (e.g., that illustrated in Figure 24.2) go beyond the scope of the reinforcement learning models discussed here (but see also Niv *et al.*, 2007).

In summary, we have suggested that goal and habit values may correspond to two computational approaches to the problem of predicting value for guiding action choice, and that the learning of habit values can be further subdivided into Pavlovian and instrumental sub-problems (Figure 24.4a).

NEURAL BASIS OF REINFORCEMENT LEARNING

The dissociation between multiple valuation mechanisms is also evident in their underlying neural substrates. Here, we review parallel evidence from animals and humans pointing to particular and distinct neural circuits supporting different forms of valuation (Figure 24.4b).

Expected Reward: Value Signals and Pavlovian Values

As described above, the reinforcement learning counterpart for Pavlovian values is the state value function. Studies of the neural basis of state or Pavlovian values have focused on three brain regions in particular: (1) the amygdala in the medial temporal lobes, (2) the orbitofrontal cortex on the ventral surface of the frontal lobes, and (3) the ventral striatum in the basal ganglia.

Single-unit recording studies in both rodents and non-human primates have implicated neurons in both amygdala and orbitofrontal cortex in encoding stimulus–reward associations (Schoenbaum *et al.*, 1998; Paton *et al.*, 2006). Cue-related and anticipatory responses that relate to the monkeys' behavioral preference for the associated outcomes have also been found in orbitofrontal cortex (Tremblay and Schultz, 1999).

Although amygdala lesions generally result in mild or no impairment of acquisition of Pavlovian or indeed instrumental associations (Baxter and Murray, 2002), lesions of either the amygdala and orbitofrontal cortex or crossed unilateral lesions of both structures result in impairments in the modulation of conditioned Pavlovian responses following changes in the value of the associated outcome, induced by an outcome devaluation procedure (e.g., pairing the outcome with a malaise-inducing substance such as lithium chloride, or feeding the animal to satiety on that outcome) (Hatfield *et al.*, 1996; Malkova *et al.*, 1997; Baxter *et al.*, 2000; Pickens *et al.*, 2003; Ostlund and Balleine, 2007).

Another brain region implicated in Pavlovian valuation is the ventral striatum, which receives afferent input from both the amygdala and orbitofrontal cortex (Fudge *et al.*, 2002; Haber *et al.*, 2006). Activity of some neurons in the ventral striatum has, as with those in the amygdala and orbitofrontal cortex, been found to reflect expected reward in relation to the onset of a stimulus presentation, and activity of neurons in this

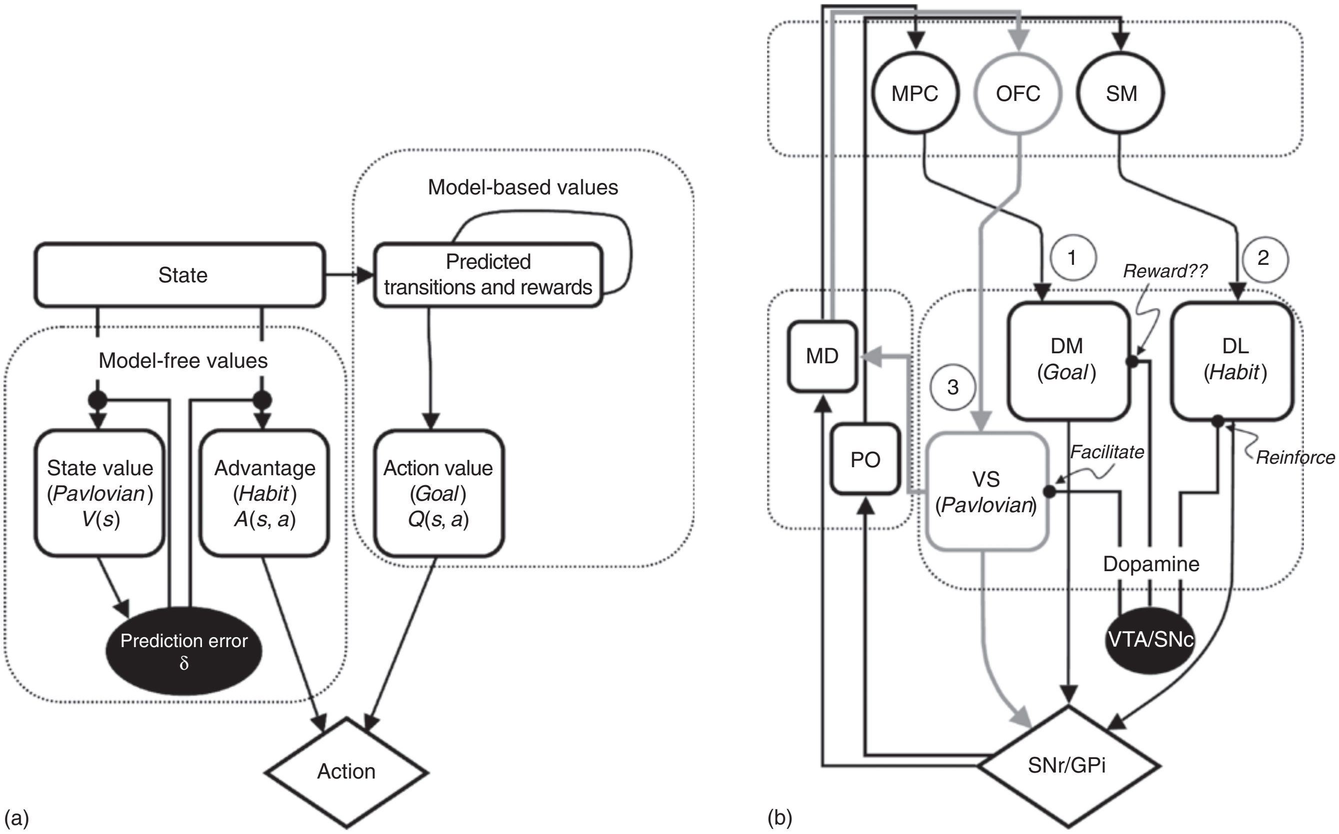


FIGURE 24.4 Computational and neural substrates of valuation. (a) Goal, habit, and Pavlovian values from model-based and model-free reinforcement learning. (b) Neural circuits underlying valuation in conditioning. Evidence reviewed in text suggests that distinct neural networks mediate the influence of goal values, habit values, and Pavlovian values on action selection and initiation. In this view, habits are encoded in a loop-like network involving sensory-motor (SM) cortical inputs to dorsolateral striatum (DL), with feedback to cortex via substantial nigra reticulata/internal segment of the globus pallidus (SNr/GPi) and posterior thalamus (PO), and are motivated by midbrain dopaminergic inputs from ventral tegmental area/substantia nigra pars compacta (VTA/SNc). A parallel circuit linking medial prefrontal cortex (MPC), dorsomedial striatum (DM), SNr, and mediodorsal thalamus (MD) mediates goal-directed actions that may speculatively involve a dopamine-mediated reward process. Finally, actions can be facilitated by Pavlovian values mediated by a parallel ventral circuit mediated by orbitofrontal cortex (OFC) and ventral striatal (VS) inputs into the habit and goal-directed loops.

area has been found to track progression through a task sequence ultimately leading to reward (Shidara *et al.*, 1998; Cromwell and Schultz, 2003; Day *et al.*, 2006). Some studies report that lesions of a part of the ventral striatum, the nucleus accumbens core, can impair learning or expression of Pavlovian approach behavior (Parkinson *et al.*, 1999). Lesions of the dopaminergic input into accumbens can also impair such learning (Parkinson *et al.*, 2002), suggesting an important role for ventral striatum in mediating learning of Pavlovian conditioned associations.

Functional neuroimaging studies in humans have also revealed activations in some of these areas during both appetitive and aversive Pavlovian conditioning in response to CS presentation (e.g. Gottfried *et al.*, 2002, 2003). When taken together, the above findings implicate a network of brain regions involving the amygdala, ventral striatum, and orbitofrontal cortex in learning and expression of Pavlovian values, which may correspond to the state value component of some reinforcement learning models.

Learning of State Value Representations

The finding of Pavlovian value signals in the brain raises the question of how such signals are learned in the first place. As we saw above, a key feature of most reinforcement learning models is the use of a prediction-error signal to update expectations of future reward based on differences between expected and actual outcomes.

Initial evidence for prediction-error signals in the brain emerged from the work of Wolfram Schultz and colleagues, who observed such signals by recording the phasic activity of dopamine neurons in awake behaving

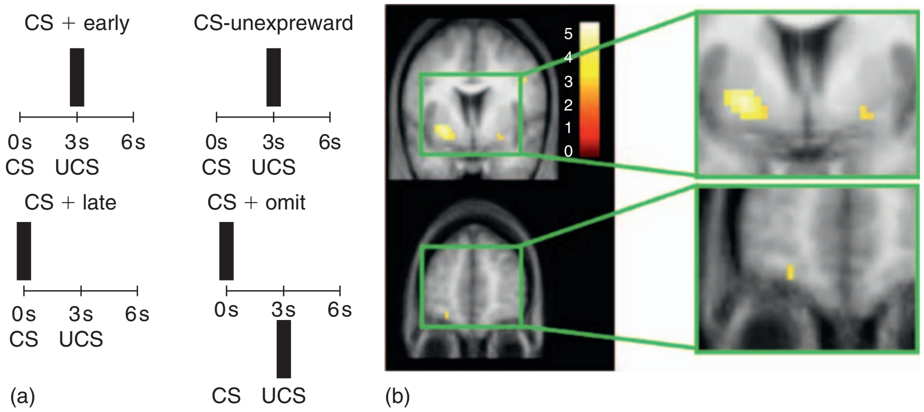


FIGURE 24.5 Prediction error signals in the human brain during appetitive Pavlovian conditioning. (a) Illustration of model prediction-error signals elicited in an experiment in which in one trial type (CS+) an arbitrary visual cue is associated 3s later with delivery of a taste reward (1-M glucose), and in another trial type (CS) a different cue is followed by no taste. In addition, occasional surprise trials occur in which the CS+ is presented but the reward is omitted (CS + omit), and the CS is presented but a reward is unexpectedly delivered (CS-unexpreward). During early CS+ trials (before learning is established) the PE signal should occur at the time of delivery of the reward, whereas by late CS+ trials (post-learning) the signal should have switched to the time of presentation of the CS. On CS + omit trials, a positive PE signal should occur at the time of presentation of the CS but a negative PE signal should occur at the time the reward was expected (CS + omit). CS-unexpreward trials should be associated with a positive signal at the time the reward is presented. (b) Parts of human ventral striatum (top) and orbitofrontal cortex (bottom) showing a significant correlation with the temporal difference prediction error signal. Data from O'Doherty *et al*. (2003).

non-human primates undergoing simple Pavlovian or instrumental conditioning tasks with reward (Mirenowicz and Schultz, 1994; Schultz *et al*., 1997; Hollerman and Schultz, 1998; Schultz, 1998). These neurons, particularly those present in the ventral tegmental area in the midbrain, project strongly to the ventral striatum, amygdala, and orbitofrontal cortex (Oades and Halliday, 1987), the three regions most closely identified as playing a role in encoding representations of stimulus-bound reward expectancies discussed earlier. The response profile of these dopamine neurons closely resembles a specific form of prediction error derived from the temporal-difference learning rule (equation 24.7), in which predictions of future reward are computed at each time interval within a trial, and the error signal is generated by computing the difference in successive predictions (Montague *et al*., 1996; Schultz *et al*., 1997). Just like the temporal-difference prediction-error signal, these neurons increase their firing when a reward is presented unexpectedly, decrease their firing from baseline when a reward is unexpectedly omitted, and respond initially at the time of the US before learning is established but shift back in time within a trial to respond instead at the time of presentation of the CS once learning has taken place (for further details, see Chapter 21 of this volume).

In order to test for evidence of a temporal-difference prediction-error signal in the human brain, O'Doherty *et al*. (2003) scanned human subjects while they underwent a classical conditioning paradigm in which associations were learned between visual stimuli and a pleasant sweet-taste reward. Significant correlations were found between a temporal-difference prediction error signal and BOLD responses in a number of brain regions, most notably the ventral striatum (ventral putamen bilaterally) (Figure 24.5) and orbitofrontal cortex, both prominent target regions of the midbrain dopamine neurons believed to carry a reward-prediction error. These results suggest that prediction-error signals are present in the human brain during Pavlovian reward-learning.

The evidence discussed so far supports the presence of prediction-error signals during learning involving appetitive or rewarding stimuli. The next obvious question is whether such signals can be found to underlie learning about punishing as well as rewarding events. Evidence of a role for dopamine neurons in responding during aversive learning is mixed. Single-unit studies have generally failed to observe strong dopaminergic activity in response to aversive events (Schultz, 1998), and indeed it has been found that dopamine neurons may in fact be inhibited in responding during aversive stimulation such as a tail-pinch in rats (Ungless *et al*., 2004). On the other hand, a number of studies measuring dopamine release in the striatum in rats using microdialysis techniques have found evidence for increased dopamine levels during aversive as well as appetitive conditioning (Pezze and Feldon, 2004). However, as termination of an aversive stimulus can in itself be rewarding, the implications of these studies for understanding the role of dopamine in aversive learning are still debated. Irrespective of whether dopamine will turn out to play a role in aversive learning or not, existing evidence appears to rule out the suggestion that phasic dopamine encodes prediction errors for aversive events in the same way that it does for appetitive events. This then raises the question of whether prediction-error signals for aversive learning are present anywhere else in the brain, such in another neuromodulatory system. The proposal has been made that serotonin neurons in the dorsal raphe might fulfill this role (Daw *et al.*, 2002).

In fact, neuroimaging studies have revealed strong evidence for the presence of prediction-error signals in the human brain during aversive as well as appetitive learning (Seymour *et al.*, 2004, 2005). Given that such aversive signals in striatum are unlikely to depend on the afferent input of dopamine neurons, these findings also show that BOLD activity in ventral striatum should not be considered to be a pure reflection of the afferent input of dopamine neurons – an interpretation implicitly assumed in some of the reward imaging literature. Rather, activity in striatum is likely to also reflect the influence of a number of different neuromodulatory systems in addition to dopamine, input from other cortical and subcortical areas, as well as intrinsic computations within this region.

The studies discussed above demonstrate that prediction-error signals are present during learning to predict both appetitive and aversive events, a finding consistent with the tenets of a prediction-error based account of associative learning. However, merely demonstrating the presence of such signals in the striatum during learning does not establish whether these signals are causally related to learning, or merely an epiphenomenon. The first study aiming to uncover a causal link was that of Pessiglione *et al.* (2006), who manipulated systemic dopamine levels by delivering a dopamine agonist and antagonist while subjects were being scanned with fMRI during performance of a reward-learning task. Prediction-error signals in striatum were boosted following administration of the dopaminergic agonist, and diminished following administration of the dopaminergic antagonist. Moreover, behavioral performance followed the changes in striatal activity and was increased following administration of the dopamine agonist and decreased following administration of the antagonist. These findings, therefore, support a causal link between prediction-error activity in striatum and the degree of behavioral learning for reward.

The Actor/Critic in the Brain

Next we turn to the mechanisms underlying action selection for reward in the brain. As outlined in above, in order to drive action selection, reinforcement learning models need to learn about the expected future reward that follows from taking individual actions, in order to use these learned action values to guide choice. The key question then, is where and how are action-related values represented in the brain? We will consider evidence to suggest the possibility that different types of action-related value representations may be encoded in different brain systems, depending on the nature of the associative relationship on which that action value depends – i.e. whether the action selection is being controlled by the goal-directed or the habitual system.

First, we will consider habit values. As outlined earlier in the chapter, these action values can be learned by model-free reinforcement learning algorithms such as the actor/critic (Barto, 1995), in which prediction-error signals generated by a critic that evaluates the expected reward available in each state of the world are used to update the policy (or stimulus–response associations) stored in a separate actor module. Thus, the implementation of an actor/critic like architecture in the brain would suggest the existence of two separate modules; a critic concerned with learning about reward expectations more generally, and an actor, which stores the action values and/or policy. Given the connections and anatomy of the basal ganglia, it has been proposed that an actor/critic-like process might be implemented in the striatum. Houk *et al.* (1995) suggested that the actor and critic could be implemented within the patch/striosome and matrix compartments distributed throughout the striatum. However, more relevant to the present discussion was the proposal by Montague *et al.* (1996) that the ventral and dorsal striatum correspond to the critic and actor respectively. The evidence reviewed above would appear to implicate a circuit involving the ventral striatum alongside the orbitofrontal cortex and amygdala in implementing a function rather similar to that proposed by the critic and involving the encoding of state values.

Evidence is also beginning to accumulate implicating at least a part of the dorsal striatum in implementing the actor. More specifically, lesions of the dorsolateral striatum have been found to abolish the onset of habitual behavior in rats, rendering the lesioned animals permanently goal-directed (Yin *et al.*, 2004). This study therefore suggests a critical involvement of the dorsolateral striatum in implementing stimulus–response learning in the rodent brain. O'Doherty *et al.* (2004) used fMRI to compare responses in an instrumental choice task to that elicited during a control Pavlovian task in which subjects experienced the same stimulus-reward contingencies but did not actively choose which action to select. This comparison was designed to isolate the actor,

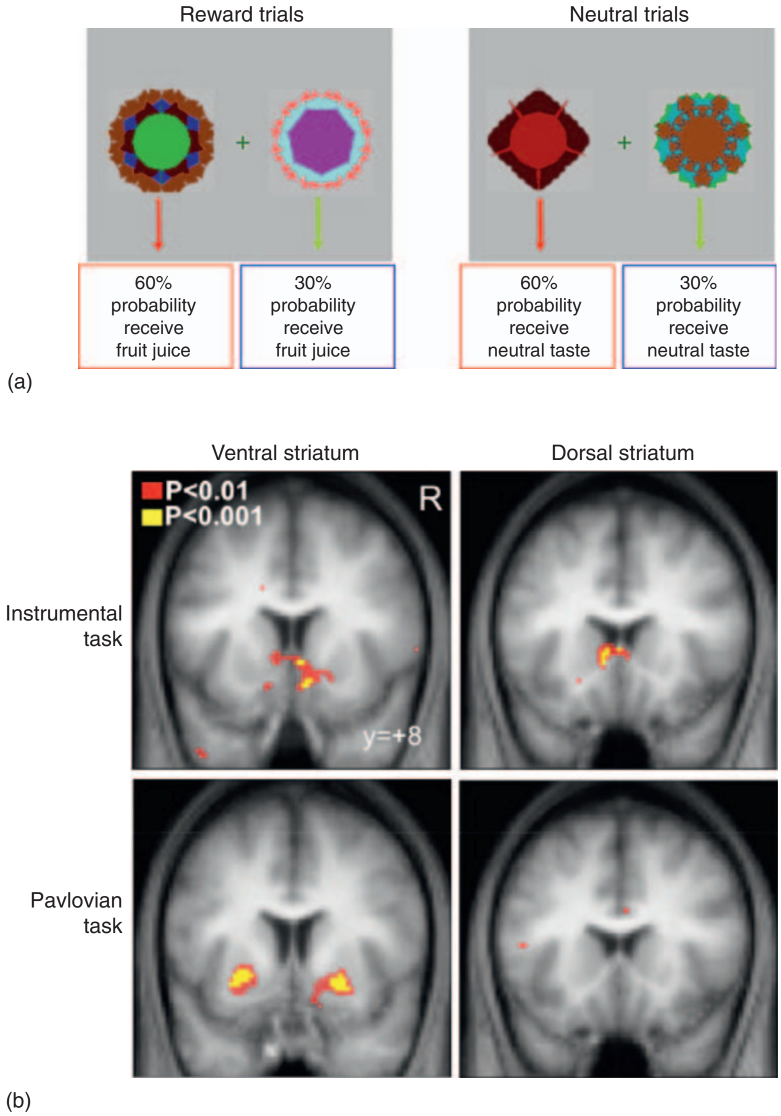


FIGURE 24.6 The actor/critic in the human striatum. (a) Instrumental choice task used by O'Doherty *et al.* (2004). On each trial of the reward condition, the subject chooses between two possible actions, one associated with a high probability (60%) and the other with a low probability (30%) of obtaining a juice reward. In a neutral condition, subjects also choose between actions with similar probabilities, but in this case they receive an affectively neutral outcome (a tasteless solution). Prediction-error responses during the reward condition of the instrumental choice task were compared to prediction-error signals during a yoked Pavlovian control task. (b) Significant correlations with the reward-prediction error signal generated by an actor/critic model were found in ventral striatum (ventral putamen extending into nucleus accumbens proper) in both the Pavlovian and instrumental tasks, suggesting that this region is involved in stimulus-outcome learning. By contrast, a region of dorsal striatum (anteromedial caudate nucleus) was found to be correlated with prediction-error signals only during the instrumental task, suggesting that this area is involved in the habitual aspects of instrumental learning. Data from O'Doherty *et al.* (2004).

which was hypothesized to be engaged only in the instrumental task, from the critic, which was hypothesized to be engaged in both the instrumental and the Pavlovian control tasks. Consistent with the proposal of a dorsal vs ventral actor/critic architecture, activity in dorsal striatum was found to be specifically correlated with prediction-error signals in the instrumental task but not in the Pavlovian task, whereas ventral striatum was correlated with prediction-error signals in both tasks (Figure 24.6) – a finding that has received support from a number of other fMRI studies (see, for example, Haruno *et al.*, 2004; Tricomi *et al.*, 2004).

It is important to note that although the architecture within the ventral and dorsal striatum is consistent with an Actor/Critic-like process, it is not yet clear whether learning of habit values in the dorsal striatum is governed by state value prediction errors generated from the critic, as should be the case for the *literal* *implementation* of the full actor/critic model. It is also plausible that habit values within the dorsal striatum might be learned and updated directly, as instantiated in other variants of reinforcement learning, such as Q-learning or SARSA (Watkins and Dayan, 1992; Sutton and Barto, 1998).

Brain Systems Underlying Goal-directed Learning in Rats and Primates

In addition to the evidence reviewed above implicating dorsolateral striatum in habitual aspects of instrumental responding, evidence is emerging that implicates a different set of brain regions in mediating the goal-directed component of instrumental responding. In rodents, two brain regions in particular have been implicated in goal-directed learning: the prelimbic cortex in the frontal lobe, and the area of striatum to which this region of cortex projects – the dorsomedial striatum. Lesions of either of these regions in rats prevent the acquisition of goal-directed learning, rendering animals habitual even during the early stages of training (Balleine and Dickinson, 1998; Corbit and Balleine, 2003; Yin *et al*., 2005). Notably, prelimbic cortex, although necessary for initial acquisition, does not appear to be necessary for the expression of goal-directed behavior, as lesions of this area do not impair goal-directed behavior once the lesions are induced after initial training (Ostlund and Balleine, 2005). On the other hand, dorsomedial striatum does appear to be critical for both learning and expression of goal-directed behavior, as lesions of this area impair such behavior if induced either before or after training (Yin *et al*., 2005).

This raises the question of whether there exist homologous regions of the primate prefrontal cortex that contribute to similar functions. A number of fMRI studies in humans have implicated parts of the frontal cortex, especially its ventral aspects, in encoding the value of chosen actions (Daw *et al*., 2006; Kim *et al*., 2006). Taken together, these findings suggest that orbital and medial prefrontal cortex are involved in keeping track of the expected future reward associated with chosen actions, and that these areas show a response profile consistent with an expected value signal generated by reinforcement learning models. However, the above studies do not differentiate between action-related value representations based on stimulus–response associations and those based on action–outcome associations.

To determine whether such signals reflect goal values or habit values, Valentin *et al*. (2007) performed an outcome devaluation paradigm in humans, similar to the manipulation used in rodents to determine whether instrumental behavior is goal-directed or habitual. By testing for regions of the brain showing a change in activity during selection of an action associated with a devalued outcome (devalued by feeding subjects to satiety on that outcome), it was possible to test for regions showing sensitivity to the learned action–outcome associations. The regions found to show such a response profile were medial and central OFC (Figure 24.7). These findings suggest that action-outcome information is present in OFC alongside stimulus-outcome representations, indicative of a role for OFC, particularly its medial aspects, in encoding expectations of reward tied to specific actions above and beyond its role in encoding stimulus-bound predictions. However, in apparent contradiction to the above findings, it has been found that lesions of OFC in rats do not produce impairments at goal-directed learning, in contrast to the effects of lesions of the prelimbic area, which do produce robust deficits in this capacity (Ostlund and Balleine, 2007).

The source of such discrepancies between studies remains to be determined, but one intriguing possibility is that rat and human OFC may not be entirely homologous. It is interesting to note that, in a previous human stimulus-based devaluation study by Gottfried *et al*. (2003), modulatory effects of reinforcer devaluation were found in central but *not* medial OFC areas, whereas in the Valentin *et al*. study, evidence was found of instrumental devaluation effects in both central *and* medial areas. This raises the possibility that medial OFC may be more involved in the goal-directed component of instrumental conditioning by encoding goal values, whereas central OFC may contribute more to Pavlovian stimulus–outcome learning by encoding Pavlovian values (as this area was found in both the Valentin *et al*. study and the previous Pavlovian devaluation study). This speculation is consistent with the known anatomical connectivity of these areas in which central areas of OFC (Brodmann areas 11 and 13) receive input primarily from sensory areas, consistent with a role for these areas in learning the relationship between stimuli, whereas the medial OFC (areas 14 and 25) receives input primarily from structures on the adjacent medial wall of prefrontal cortex, such as cingulate cortex – an area often implicated in response selection and/or reward-based action choice (Carmichael and Price, 1995, 1996). It is also notable that although the majority of single-unit studies in monkeys have reported stimulus-related activity and not response-related selectivity in the OFC (e.g., Thorpe *et al*., 1983; Tremblay and Schultz, 1999; Padoa-Schioppa and Assad, 2006), these studies have typically recorded from more lateral and

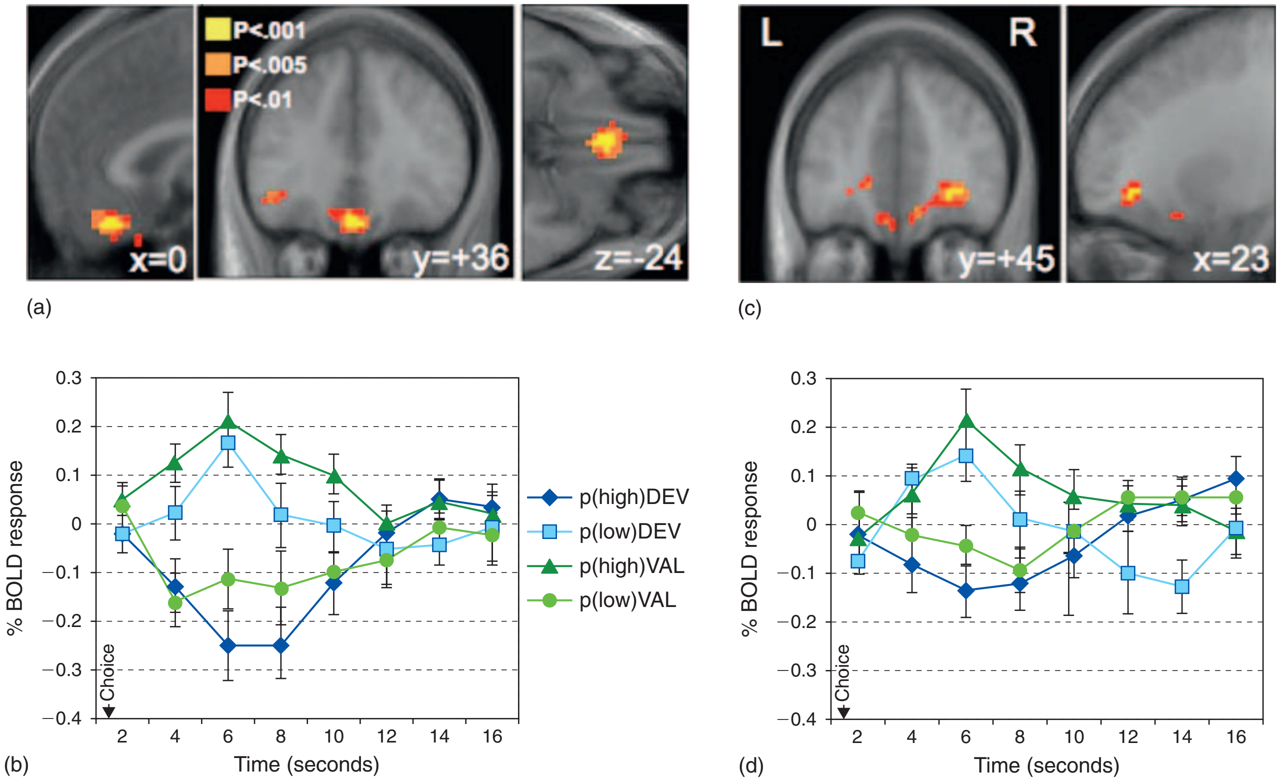


FIGURE 24.7 Regions of human OFC exhibiting response properties consistent with action-outcome learning. Neural activity in OFC during action selection for reward, showing a change in response properties as a function of the value of the outcome with each action. Choice of an action leading to a high probability of obtaining an outcome that had been devalued (*P*(high)dev) led to a decrease in activity in these areas, whereas choice of an action leading to a high probability of obtaining an outcome that was still valued led to an increase in activity in the same areas. Devaluation was accomplished by means of feeding the subject to satiety on that outcome prior to the test period. (a) A region of medial OFC showing a significant modulation in its activity during instrumental action selection as a function of the value of the associated outcome. (b) Time-course plots derived from the peak voxel (from each individual subject) in the mOFC during trials in which subjects chose each one of the four different actions (choice of the high- vs low-probability action in either the valued or the devalued conditions). (c) A region of right central OFC also showing a significant interaction effect. (d) Time-course plots from the peak voxel (from each individual subject) in the right central OFC. Data from Valentin *et al*. (2007), with permission; ©The Society for Neuroscience.

central areas of the OFC (Brodmann areas 12/47 and 13, respectively), and not from more medial areas. It is therefore plausible that more medial sectors of the OFC in humans correspond to regions considered part of medial prefrontal cortex in rats, and that have been more conclusively linked to goal-directed learning (Corbit and Balleine, 2003; Balleine and Dickinson, 1998; Ostlund and Balleine, 2005). Indeed, consistent with this hypothesis, Ostlund and Balleine (2007) found that whereas medial prefrontal cortex was involved in goal-directed action, lesion-induced damage to the more central and lateral regions of the OFC, while leaving goal-directed action intact, abolished the ability of Pavlovian cues to selectively elevate instrumental performance in an outcome-selective Pavlovian-instrumental transfer task (such as that illustrated in Figure 24.2).

As outlined earlier in the chapter, a computational framework for goal-directed learning is to propose that this form of learning is mediated by a form of "model-based" inference in which the agent uses a full model of the decision problem to iteratively evaluate the future consequences of each action in order to compute action values. Evidence that model-based inference may be implemented in the human ventromedial prefrontal cortex has emerged from a study by Hampton *et al*. (2006), wherein subjects were scanned with fMRI while participating in a decision problem called *probabilistic reversal learning*, which has a hidden structure in that rewards available following choice of two different actions are anti-correlated, and the values of the two actions reverse from time to time. Hampton and colleagues compared two computational models in terms of how well they could account for human choice behavior in the task and for the pattern of neural activity in vmPFC during task performance. One of these algorithms incorporated the rules or structure of the decision problem as would be expected for a model-based inference mechanism, whereas the other model was a simple model-free reinforcement learning algorithm that did not incorporate the structure and thus would only learn to slowly and incrementally update values following successive reinforcements. Consistent with a role for vmPFC in model-based inference, predicted reward signals in this region were found to reflect the structure of the decision problem, such that activity was updated instantly following a reversal rather than being updated incrementally, as might be expected for a model-free reinforcement learning mechanism.

CONCLUSIONS

We have reviewed evidence that the values underpinning decision making are not unitary, but instead are fractionated psychologically, neurally, and computationally. In particular, we have distinguished goal values from habit values: the former are more cognitive and grounded in knowledge about particular expected outcomes, while the latter reflect more generalized motivation divorced from any particular goal identity. Computationally, this distinction parallels one between different methods for bringing experience to bear on the problem of learned value predictions. The influences of these values on behavior are dissociable through manipulations such as reward devaluation, which affects goal values but not habit values. Neurally, lesions localize the two sorts of value in discrete cortico-striatal networks, comprising medial vs lateral portions of dorsal striatum, together with their associated areas of prefrontal cortex and thalamus. Candidate human analogues of many parts of these networks have been identified using functional neuroimaging.

We have also identified a third sort of value – the Pavlovian value associated with stimuli or states. Though such values are not associated with actions, they nonetheless can impact behavior in a number of ways; most particularly these values can affect choice, apparently by providing information as to the likelihood that an action will pay off with a specific outcome or consequence. Yet a third region of striatum, the ventral part, is an important locus for Pavlovian values and their influences on instrumental behavior.

Two major issues remain. First, what role does dopamine play in each of these systems? Computational models such as the actor/critic identify dopamine in ventral and dorsolateral striatum as a signal for training Pavlovian and habitual values – a role the system is anatomically well situated and physiologically well suited to play; particularly given that, according to the models, a single prediction-error signal can train both sorts of value. However, the involvement of dorsomedial striatum in goal values raises the likelihood of dopaminergic participation in these as well, since this area's tight interrelationship with dopaminergic midbrain parallels that of the other striatal subterritories (Calabresi *et al.*, 2007). From the perspective of computational models of goal-directed action, this is a puzzling prospect, because the learning underlying goal values is envisioned to be of a rather different sort, and because characteristic features of the dopaminergic response (such as the transfer of responses from rewards to stimuli predicting them) seem closely related to schemes (such as "caching"; Daw *et al.*, 2005) for learning habit values. Clearly, more data are required regarding dopamine's involvement with goal values. The single most crucial question is, how dopamine neurons behave following reward devaluation? Does the dopaminergic response to a state or action persist when its associated reward is devalued, as would be expected if the response were driven by habit values? Any indications of devaluation sensitivity or outcome specificity in the dopamine response would motivate a rethinking of computational accounts of dopamine, and an investigation of hybrid models (like the successor representation; Dayan, 1993) that use habit-style methods to learn outcome-specific goal values.

A second open issue is to what extent our framework of dissociable value systems, whose roots lie largely in animal conditioning, overlaps other dual- or multi-system models popular in human decision making. Such models are particularly common in behavioral economic approaches to self-control issues such as temporal discounting (Thaler, 1981; Loewenstein and Prelec, 1992). To develop one particular example, several models (see, for example, Laibson, 1997) hold that inter-temporal choice is mediated by two competing systems; an impulsive system in which future rewards are highly discounted, and a reflective system which discounts future rewards only shallowly. While an obvious possibility is to identify these influences with habitual and goal-directed systems, respectively, it has also been suggested that Pavlovian values might produce impulsive effects in many time-discounting paradigms (Dayan *et al.*, 2006). Neurally, McClure *et al.* (2004) have proposed, on the basis of an fMRI study of temporal discounting behavior, that a reflective system is localized in lateral parts of prefrontal cortex, whereas an impulsive system is suggested to be present in the striatum. The evidence behind this putative dissociation has been hotly disputed (Kable and Glimcher, 2007). In any case, this proposed anatomical breakdown does not map perfectly onto the goal-directed vs habitual framework outlined here, because, as discussed above, different parts of the dorsal striatum are involved in both goal-directed and habitual processes. Also, though the goal-directed system, like the reflective system, is suggested to depend on prefrontal cortex, our imaging and lesion work have implicated medial rather than lateral aspects of these cortices. An important area for future research will be to categorize similarities and differences between these different possible models of human choice, in order ultimately to develop a more parsimonious and unitary model that accounts for each of the different behaviors and most closely reflects the true structure (both computationally and neuro-anatomically) of the brain's systems for guiding choice.

References

Adams, C.D. (1981). Variations in the sensitivity of instrumental responding to reinforcer devalaution. *Q. J. Exp. Psychol.* 34B, 77–98.

Adams, C.D. and Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Q. J. Exp. Psychol.* 33B, 109–121.

Antonini, A., Moresco, R.M., Gobbo, C. *et al*. (2001). The status of dopamine nerve terminals in Parkinson's disease and essential tremor: a PET study with the tracer [11-C]FE-CIT. *Neurol. Sci.* 22, 47–48.

Baird, L.C. (1994). Reinforcement learning in continuous time: advantage updating. In: *Proceedings of the International Conference on Neural Networks* 4, 2448–2453.

Balleine, B.W. (2001). Incentive processes in instrumental conditioning. In: R.M.S. Klein (ed.), *Handbook of Contemporary Learning Theories*. Hillsdale, NJ: LEA, pp. 307–366.

Balleine, B.W. and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419.

Balleine, B.W. and Ostlund, S.B. (2007). Still at the choice point: action selection and initiation in instrumental conditioning. *Ann. NY Acad. Sci.* 1104, 147–171.

Barto, A.G. (1992). Reinforcement learning and adaptive critic methods. In: D.A. White and D.A. Sofge (eds), *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches*. New York, NY: Van Norstrand Reinhold, pp. 469–491.

Barto, A.G. (1995). Adaptive critics and the basal ganglia. In: J.C. Houk, J.L. Davis, and D.G. Beiser (eds), *Models of Information Processing in the Basal Ganglia*. Cambridge, MA: MIT Press, pp. 215–232.

Baxter, M.G. and Murray, E.A. (2002). The amygdala and reward. *Nat. Rev. Neurosci.* 3, 563–573.

Baxter, M.G., Parker, A., Lindner, C.C. *et al*. (2000). Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *J. Neurosci.* 20, 4311–4319.

Bloch, M.H., Leckman, J.F., Zhu, H., and Peterson, B.S. (2005). Caudate volumes in childhood predict symptom severity in adults with Tourette syndrome. *Neurology* 65, 1253–1258.

Calabresi, P., Picconi, B., Tozzi, A., and Di Filippo, M. (2007). Dopamine-mediated regulation of corticostriatal synaptic plasticity. *Trends Neurosci.* 30, 211–219.

Cardinal, R.N. and Everitt, B.J. (2004). Neural and psychological mechanisms underlying appetitive learning: links to drug addiction. *Curr. Opin. Neurobiol.* 14, 156–162.

Carmichael, S.T. and Price, J.L. (1995). Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *J. Comp. Neurol.* 363, 642–664.

Carmichael, S.T. and Price, J.L. (1996). Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *J. Comp. Neurol.* 371, 179–207.

Chang, J.Y., Chen, L., Luo, F. *et al*. (2002). Neuronal responses in the frontal cortico-basal ganglia system during delayed matching-to-sample task: ensemble recording in freely moving rats. *Exp. Brain Res.* 142, 67–80.

Colwill, R.C. and Rescorla, R.A. (1986). Associative structures in instrumental learning. *Psychol. Learning Motiv.* 20, 55–104.

Colwill, R.M. and Rescorla, R.A. (1988). Associations between the discriminative stimulus and the reinforcer in instrumental learning. *J. Exp. Psychol. Animal Behav. Proc.* 14, 155–164.

Colwill, R.M. and Motzkin, D.K. (1994). Encoding of the unconditioned stimulus in Pavlovian conditioning. *Animal Learning Behav.* 22, 384–394.

Corbit, L.H. and Balleine, B.W. (2003). The role of prelimbic cortex in instrumental conditioning. *Behav. Brain Res.* 146, 145–157.

Corbit, L.H. and Balleine, B.W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of Pavlovian-instrumental transfer. *J. Neurosci.* 25, 962–970.

Corbit, L.H., Muir, J.L., and Balleine, B.W. (2001). The role of the nucleus accumbens in instrumental conditioning: evidence of a functional dissociation between accumbens core and shell. *J. Neurosci.* 21, 3251–3260.

Cromwell, H.C. and Schultz, W. (2003). Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J. Neurophysiol.* 89, 2823–2838.

Davis, J. and Bitterman, M.E. (1971). Differential reinforcement of other behavior (DRO): a yoked-control comparision. *J. Exp. Anal. Behav.* 15, 237–241.

Daw, N.D., Kakade, S., and Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Networks* 15, 603–616.

Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and striatal systems for behavioral control. *Nature Neurosci.* 8, 1704–1711.

Daw, N.D., O'Doherty, J.P., Dayan, P. *et al*. (2006). Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.

Day, J.J., Wheeler, R.A., Roitman, M.F., and Carelli, R.M. (2006). Nucleus accumbens neurons encode Pavlovian approach behaviors: evidence from an autoshaping paradigm. *Eur. J. Neurosci.* 23, 1341–1351.

Dayan, PO. (1993). Improving generalisation for temporal difference learning: the successor representation. *Neural Computation* 5, 613–624.

Dayan, P. and Balleine, B.W. (2002). Reward, motivation and reinforcement learning. *Neuron* 36, 285–298.

Dayan, P., Niv, Y., Seymour, B., and Daw, N.D. (2006). The misbehavior of value and the discipline of the will. *Neural Networks* 19, 1153–1160.

Delamater, A.R. (1995). Outcome-selective effects of intertrial reinforcement in Pavlovian appetitive conditioning with rats. *Animal Learning Behav.* 23, 31–39.

Dickinson, A. (1994). Instrumental conditioning. In: N.J. Mackintosh (ed.), *Animal Cognition and Learning*. London: Academic Press, pp. 4–79.

Dickinson, A. and Balleine, B.W. (1994). Motivational control of goal-directed action. *Animal Learning Behav.* 22, 1–18.

Dickinson, A. and Mulatero, C.W. (1989). Reinforcer specificity of the suppression of instrumental performance on a non-contingent schedule. *Behav. Processes* 19, 167–180.

Dickinson, A., Nicholas, D.J., and Adams, C.D. (1983). The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Q. J. Exp. Psychol.* 35B, 35–51.

Dickinson, A., Balleine, B.W., Watt, A. *et al*. (1995). Overtraining and the motivational control of instrumental action. *Animal Learning Behav.* 22, 197–206.

Dickinson, A., Squire, S., Varga, Z., and Smith, J.W.. (1998). Omission learning after instrumental pretraining. *Q. J. Exp. Psychol.* 51, 271–286.

Dickinson, A., Wood, N., and Smith, J.W. (2002). Alcohol seeking by rats: action or habit? *Q. J. Exp. Psychol. B* 55, 331–348.

Frankland, P.W., Wang, Y., Rosner, B. *et al*. (2004). Sensory-gating abnormalities in young males with fragile X syndrome and Fmr1-knockout mice. *Mol. Psych.* 9, 417–425.

Fudge, J.L., Kunishio, K., Walsh, P. *et al*. (2002). Amygdaloid projections to ventromedial striatal subterritories in the primate. *Neuroscience* 110, 257–275.

Fuster, J.M. (2000). Executive frontal functions. *Exp. Brain Res.* 133, 66–70.

Goldman-Rakic, P.S. (1995). Architecture of the prefrontal cortex and the central executive. *Ann. NY Acad. Sci.* 769, 71–83.

Gottfried, J.A., O'Doherty, J., and Dolan, R.J. (2002). Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. *J. Neurosci.* 22, 10829–10837.

Gottfried, J.A., O'Doherty, J., and Dolan, R.J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* 301, 1104–1107.

Haber, S.N., Kim, K.S., Mailly, P., and Calzavara, R. (2006). Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. *J. Neurosci.* 26, 8368–8376.

Hampton, A.N., Bossaerts, P., and O'Doherty, J.P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J. Neurosci.* 26, 8360–8367.

Haruno, M. and Kawato, M. (2006). Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops: fMRI examination in stimulus–action–reward association learning. *Neural Networks* 19, 1242–1254.

Haruno, M., Kuroda, T., Doya, K. *et al*. (2004). A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. *J. Neurosci.* 24, 1660–1665.

Hatfield, T., Han, J.S., Conley, M. *et al*. (1996). Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. *J. Neurosci.* 16, 5256–5265.

Hodges, A., Strand, A.D., Aragaki, A.K. *et al*. (2006). Regional and cellular gene expression changes in human Huntington's disease brain. *Hum. Mol. Genet.* 15, 965–977.

Holland, P.C. (1979). Differential effects of omission contingencies on various components of Pavlovian appetitive conditioned responding in rats. *J. Exp. Psychol. Animal Behav Proc.* 5, 178–193.

Holland, P.C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J. Exp. Psychol. Animal Behav. Process* 30, 104–117.

Holland, P.C. and Gallagher, M. (2004). Amygdala–frontal interactions and reward expectancy. *Curr. Opin. Neurobiol.* 14, 148–155.

Hollerman, J.R. and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309.

Holman, E.W. (1975). Some conditions for the dissociation of consummatory and instrumental behavior in rats. *Learning Motiv.* 6, 358–366.

Houk, J.C., Adams, J.L., and Barto, A.G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: J.C. Houk, J.L. Davis, and B.G. Beiser (eds), *Models of Information Processing in the Basal Ganglia*. Cambridge, MA: MIT Press, pp. 249–270.

Hull, C.L. (1943). *Principles of Behavior*. New York, NY: Appleton.

Jog, M.S., Kubota, Y., Connolly, C.I. *et al*. (1999). Building neural representations of habits. *Science* 286, 1745–1749.

Kable, J.W. and Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.

Kim, H., Shimojo, S., and O'Doherty, J.P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biol.* 4, e233.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Q. J. Economics* 112, 443–477.

Lauwereyns, J., Watanabe, K., Coe, B., and Hikosaka, O. (2002). A neural correlate of response bias in monkey caudate nucleus. *Nature* 418, 413–417.

Levy, R. and Dubois, B. (2006). Apathy and the functional anatomy of the prefrontal cortex–basal ganglia circuits. *Cerebral Cortex* 16, 916–928.

Loewenstein, G. and Prelec, D. (1992). Anomalies in intertemporal choice: evidence and an interpretation. *Q. J. Economics* 107, 573–597.

Malkova, L., Gaffan, D., and Murray, E.A. (1997). Excitotoxic lesions of the amygdala fail to produce impairment in visual learning for auditory secondary reinforcement but interfere with reinforcer devaluation effects in rhesus monkeys. *J. Neurosci.* 17, 6011–6020.

McClure, S.M., Laibson, D.I., Loewenstein, G., and Cohen, J.D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507.

Miles, F.J., Everitt, B.J., and Dickinson, A. (2003). Oral cocaine seeking by rats: action or habit? *Behav. Neurosci.* 117, 927–938.

Mirenowicz, J. and Schultz, W. (1994). Importance of unpredictability for reward responses in primate dopamine neurons. *J. Neurophysiol.* 72, 1024–1027.

Montague, P.R., Dayan, P., and Sejnowski, T.J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.

Niv, Y., Joel, D., and Dayan, P. (2006). A normative perspective on motivation. *Trends Cogn. Sci.* 10, 375–381.

Niv, Y., Daw, N.D., Joel, D., and Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology* 191, 507–520.

Oades, R.D. and Halliday, G.M. (1987). Ventral tegmental (A10) system: neurobiology. 1. Anatomy and connectivity. *Brain Res.* 434, 117–165.

O'Doherty, J., Dayan, P., Friston, K. *et al*. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron* 38, 329–337.

O'Doherty, J., Dayan, P., Schultz, J. *et al*. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454.

Ostlund, S.B. and Balleine, B.W. (2005). Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal–directed learning. *J. Neurosci.* 25, 7763–7770.

Ostlund, S.B. and Balleine, B.W. (2007). Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. *J. Neurosci.* 27, 4819–4825.

Padoa-Schioppa, C. and Assad, J.A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.

Parkinson, J.A., Olmstead, M.C., Burns, L.H. *et al*. (1999). Dissociation in effects of lesions of the nucleus accumbens core and shell on appetitive pavlovian approach behavior and the potentiation of conditioned reinforcement and locomotor activity by D-amphetamine. *J. Neurosci.* 19, 2401–2411.

Parkinson, J.A., Dalley, J.W., Cardinal, R.N. *et al*. (2002). Nucleus accumbens dopamine depletion impairs both acquisition and performance of appetitive Pavlovian approach behaviour: implications for mesoaccumbens dopamine function. *Behav. Brain Res.* 137, 149–163.

Parkinson, J.A., Roberts, A.C., Everitt, B.J., and Di Ciano, P. (2005). Acquisition of instrumental conditioned reinforcement is resistant to the devaluation of the unconditioned stimulus. *Q. J. Exp. Psychol. B* 58, 19–30.

Partridge, J.G., Tang, K.C., and Lovinger, D.M. (2000). Regional and postnatal heterogeneity of activity-dependent long-term changes in synaptic efficacy in the dorsal striatum. *J. Neurophysiol.* 84, 1422–1429.

Paton, J.J., Belova, M.A., Morrison, S.E., and Salzman, C.D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439, 865–870.

Pavlov, I.P. (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. Oxford: Oxford University Press.

Pessiglione, M., Seymour, B., Flandin, G. *et al*. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442, 1042–1045.

Pezze, M.A. and Feldon, J. (2004). Mesolimbic dopaminergic pathways in fear conditioning. *Prog. Neurobiol.* 74, 301–320.

Pickens, C.L., Saddoris, M.P., Setlow, B. *et al*. (2003). Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *J. Neurosci.* 23, 11078–11084.

Pithers, R.T. (1985). The roles of event contingencies and reinforcement in human autoshaping and omission responding. *Learning Motiv.* 16, 210–237.

Poldrack, R.A., Clark, J., Pare-Blagoev, E.J. *et al*. (2001). Interactive memory systems in the human brain. *Nature* 414, 546–550.

Rescorla, R.A. and Solomon, R.L. (1967). Two-process learning theory: relationships between Pavlovian conditioning and instrumental learning. *Psychol. Rev.* 74, 151–182.

Robbins, T.W. and Everitt, B.J. (2002). Limbic-striatal memory systems and drug addiction. *Neurobiol. Learning Mem.* 78, 625–636.

Robinson, D., Wu, H., Munne, R.A. *et al*. (1995). Reduced caudate nucleus volume in obsessive-compulsive disorder. *Arch. Gen. Psych.* 52, 393–398.

Schoenbaum, G., Chiba, A.A., and Gallagher, M. (1998). Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat. Neurosci.* 1, 155–159.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Seppi, K., Schocke, M.F., Prennschuetz-Schuetzenau, K. *et al*. (2006). Topography of putaminal degeneration in multiple system atrophy: a diffusion magnetic resonance study. *Mov. Disord.* 21, 847–852.

Seymour, B., O'Doherty, J.P., Dayan, P. *et al*. (2004). Temporal difference models describe higher-order learning in humans. *Nature* 429, 664–667.

Seymour, B., O'Doherty, J.P., Koltzenburg, M. *et al*. (2005). Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat. Neurosci.* 8, 1234–1240.

Sheffield, F.D. (1965). Relation between classical and instrumental conditioning. In: W.F. Prokasy (ed.), *Classical Conditioning*. New York, NY: Appleton Century Crofts, pp. 302–322.

Shidara, M., Aigner, T.G., and Richmond, B.J. (1998). Neuronal signals in the monkey ventral striatum related to progress through a predictable series of trials. *J. Neurosci.* 18, 2613–2625.

Smith, R., Musleh, W., Akopian, G. *et al*. (2001). Regional differences in the expression of corticostriatal synaptic plasticity. *Neuroscience* 106, 95–101.

Sutton, R.S. and Barto, A.G. (1998). *Reinforcement Learning*. Cambridge, MA: MIT Press.

Tanaka, S.C., Samejima, K., Okada, G. *et al*. (2006). Brain mechanism of reward prediction under predictable and unpredictable environmental dynamics. *Neural Networks* 19, 1233–1241.

Thaler, R.H. (1981). Some empirical evidence on time inconsistency. *Rev. Econ. Stud.* 23, 165–180.

Thorndike, E.L. (1911). *Animal Intelligence: Experimental Studies*. New York, NY: Macmillan.

Thorpe, S.J., Rolls, E.T., and Maddison, S. (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp. Brain Res.* 49, 93–115.

Trapold, M.A. and Overmier, J.B. (1972). The second learning process in instrumental conditioning. In: A.A. Black and W.F. Prokasy (eds), *Classical Conditioning: II. Current Research and Theory*. New York, NY: Appleton Century Crofts, pp. 427–452.

Tremblay, L. and Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature* 398, 704–708.

Tricomi, E.M., Delgado, M.R., and Fiez, J.A. (2004). Modulation of caudate activity by action contingency. *Neuron* 41, 281–292.

Ungless, M.A., Magill, P.J., and Bolam, J.P. (2004). Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science* 303, 2040–2042.

Valentin, V.V., Dickinson, A., and O'Doherty, J.P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* 27, 4019–4026.

Watkins, C.J.C.H. (1989). *Learning from delayed rewards*. PhD thesis, University of Cambridge.

Watkins, C.J. and Dayan, P. (1992). Q-learning. *Machine Learning* 8, 279–292.

Williams, D.R. and Williams, H. (1969). Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. *J. Exp. Anal. Behav.* 12, 511–520.

Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* 19, 181–189.

Yin, H.H., Ostlund, S.B., Knowlton, B.J., and Balleine, B.W. (2005). The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* 22, 513–523.

CHAPTER

25

Representation of Subjective Value in the Striatum

Brian Knutson, Mauricio R. Delgado and Paul E.M. Phillips

OUTLINE

INTRODUCTION

Overview

A primary goal of neuroeconomics is to explain how brains choose. In contrast to a correlational approach, recent technological innovations facilitate a process approach that examines how distinct neural components promote subsequent choice. We review animal and human studies that focus on the representation of subjective value in the striatum. Converging evidence suggests that ventral striatal regions represent anticipated value while dorsal striatal regions represent the value of outcomes in the service of choosing future actions, suggesting a temporal flow of information through the striatum during incentive processing in the service of estimating and acquiring gains. These emerging findings highlight the dynamic, componential, and ultimately subjective nature of valuation, and hold implications for informing public policy and diagnosing mental illness.

Background

While neuroeconomics aims to explain how the brain chooses (i.e., often with respect to resource allocation), different disciplines have historically approached this question from different vantage points. Three prototypical views can be identified – those of neuroscientists, psychologists, and economists. Neuroscientists typically start from neurons, or the cells that make up the brain. Psychologists typically start from experiential phenomena related to affect, cognition, and behavior. Economists typically start from axioms or mathematically consistent specifications. These different views can subtly bias the assumptions and goals of each type of investigator. For instance, neuroscientists might want to use axioms to model neural activity, while economists might want to use neural activity to verify axioms, and psychologists might want to map intermediate links between neural activity and axioms (including affect, cognition, and behavior) (see also Chapter 10 of this volume). All researchers, however, can benefit from using neuroscientific methods to improve prediction of choice.

To investigate how the brain chooses, researchers have typically adopted one of two strategies. One "correlational" strategy seeks to identify neural correlates of choice (consistent with behavioral neuroscience, behavioral psychological, and revealed-preference economic approaches). Another "process" strategy seeks to determine how different neural components causally influence future choice (consistent with cognitive neuroscience, cognitive psychology, and prospect-theoretic economic approaches). Since a process strategy assumes that value is represented before choice, it implicitly encourages dynamic and componential decomposition of value representation, which should then facilitate prediction. At minimum, for instance, valuation can be broken down into positive and negative representations of anticipated and received incentives (Figure 25.1; Knutson and Cooper, 2005). Critically, such a decomposition does not assume that gain is the opposite of loss, or that anticipation is the same as outcome. Methods with advanced spatiotemporal resolution can facilitate process analyses by enabling researchers to measure value representation before as well as after choice. For the purposes of this chapter, we specifically refer to subjective valuation, or the value that a given individual (rather than an experimenter or society) places on a stimulus. This subjective value is computed before contact with the stimulus, and can facilitate approach towards or avoidance of the stimulus, but also can be updated based on experience with the stimulus. Thus, we minimally assume that subjective value is represented in the brain prior to choice (potentially by multiple components), can influence subsequent choice, and can be dynamically assigned (and thus updated). Consistent with these criteria, below we review (1) historical findings implicating the striatum in the representation of subjective value, (2) structural delineation of different striatal components, (3) functional studies attempting to localize subjective value to striatal components, (4) prediction studies attempting to use striatal signals to predict choice, and (5) conclusions and implications.

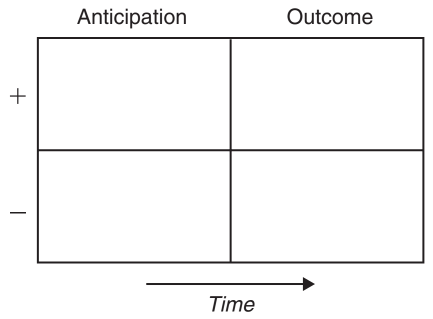


FIGURE 25.1 Minimal scheme for a process analysis of incentive valuation. Reproduced from Knutson and Cooper (2005), with permission.

History

One popular "information processing" analogy likens the brain to a computer (Newell and Simon, 1972). This analogy fits to some extent, since the brain includes many small but interconnected components (neurons) capable of converting continuous (or analog) to binary (or digital) signals (Rumelhart and McClelland, 1986). However, the analogy's fit is also somewhat forced, since it fails to consider a critical functional constraint on brain design. The brain is not like just any computer – it is a computer specifically designed to facilitate its host's survival and procreation (Dawkins, 1989; Glimcher, 2003). Thus, the brain must subjectively evaluate incoming information in the light of these overarching goals before acting upon it. Conceptually, this scheme implies an inward funneling in which sensory impulses must be filtered and then subjectively evaluated, followed by an outward funneling in which completed evaluations then flexibly link to appropriate motor responses. Physiologically, this scheme implies the existence of evolutionarily-conserved subjective valuation mechanisms that can link sensory impressions to motor responses.

Where in the brain's haystack of neurons might the needles of subjective value lie? The funneling scheme implies that if one could directly stimulate neural circuits that represent value, one could unconditionally elicit relevant behavior (i.e., approach or avoidance). Following lesions and recording, brain stimulation is one of the oldest of neurophysiological methods. While introductory neuroscience textbooks typically feature studies in which researchers used electrical stimulation to map sensory and motor cortical organization (Jasper and Penfield, 1954), fewer focus on electrical stimulation of subcortical regions. As illustrated by early studies in which subcortical electrical stimulation unconditionally elicited enraged behavior in cats (Hess, 1959), electrical stimulation of subcortical regions can unconditionally elicit approach or avoidance behavior. For instance, decades of research suggest that electrical stimulation of brain regions that lie along the ascending trajectory of dopamine neurons (i.e., projecting from the ventral tegmental area to the lateral hypothalamus to ventral striatal regions and the medial prefrontal cortex) can unconditionally elicit approach behavior (Olds and Fobes, 1981), while electrical stimulation of other regions (i.e., descending from the insular cortex and lateral amygdala to the stria terminals and medial hypothalamus to the periaqueductal gray) unconditionally elicits avoidance behavior (LeDoux, 2000). Beyond experimenter-administered electrical stimulation, animals also self-administer brain stimulation, as initially dramatically demonstrated in rats (Olds and Milner, 1954). Many of the regions that elicit self-administration overlap with those that unconditionally elicit approach behavior after stimulation (listed above), implicating these regions in motivation as well as motion. Some of the regions that elicit the most vigorous self-stimulation include ventral parts of the striatum (Olds and Fobes, 1981). Thus, striatal circuits provide a promising starting point in the search for the neural substrates of subjective value.

STRUCTURE

The striatum is a set of subcortical structures located in the center of the brain, behind the frontal lobes and encircling the thalamus. The striatum includes three structures: the caudate, the putamen, and the nucleus accumbens (NAcc). The caudate (Latin: "having a tail") is the most superior, and has the appearance of a comet in its anterior aspect, with a tail curving around both sides of the lateral ventricles towards the posterior part of the brain. Thus, parts of the caudate have been named the head, the medial aspect, and the tail, curving from front to back. The putamen (Latin: "peel, husk, or shell of seed or fruit") is the most inferior, and has the appearance of a shell, curving upward from the base of the brain. The caudate and putamen are diagonally divided in primates (but not in rodents, in which they are called the caudate-putamen) by a white matter tract called the internal capsule, which carries nerves from the motor cortex to the red nucleus near the base of the brain (Herrero *et al.*, 2002). In fact, the striatum (Latin: "striped") takes its name from the striped appearance of the internal capsule due to interposed gray (neurons) and white (glia) matter (Finger, 1994). The nucleus accumbens (Latin: "kernel which lies against" the wall of the third ventricle) has the appearance of two tubes connecting the caudate and putamen ventrally, where the internal capsule ends (Figure 25.2). While comparative researchers have divided the NAcc into medial shell and lateral core aspects using cell identification techniques (Zahm and Brog, 1992), most neuroimaging methods presently lack such fine spatial resolution.

As is common in neuroanatomical nomenclature, other terms refer to striatal structures or their parts, potentially causing confusion. The "basal ganglia" (Latin: "lower swelling") refers to a set of subcortical gray matter structures that includes the striatum, but also the globus pallidus (Latin: "pale clump," a motor output region), the subthalamic nucleus (a relay to the prefrontal cortex), and the substantia nigra (Latin: "black substance", which provides dopaminergic

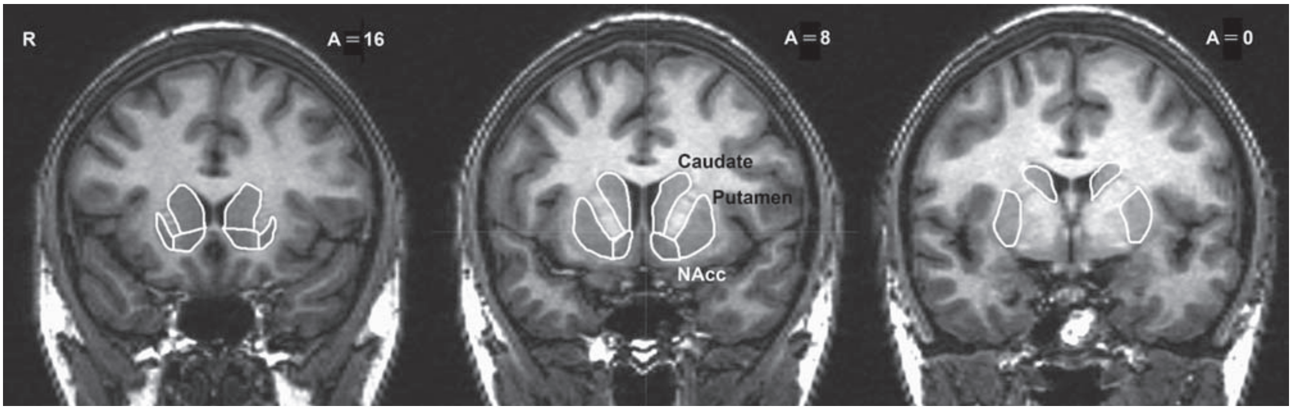


FIGURE 25.2 Nucleus accumbens (NAcc), caudate, and putamen. (as specified by Breiter *et al.*, 1997.)

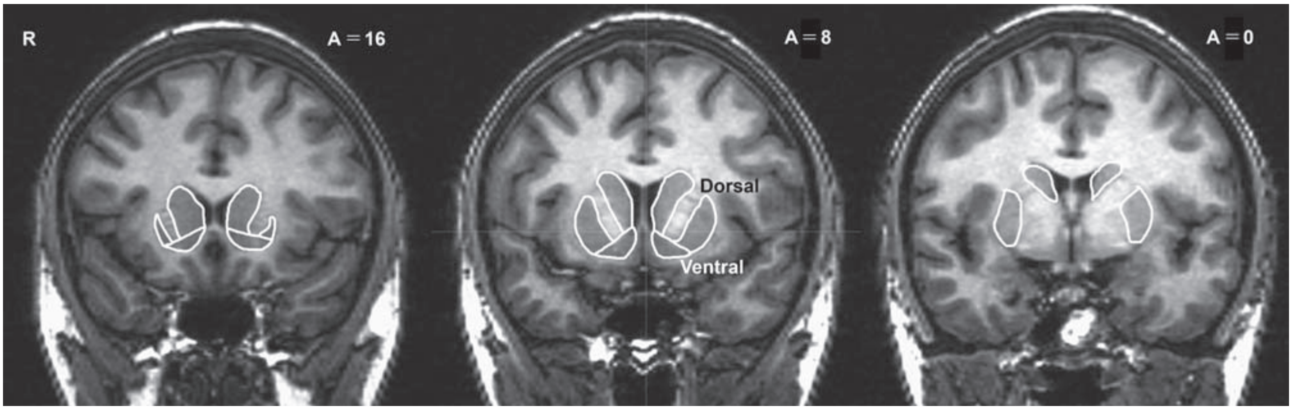


FIGURE 25.3 Ventral and dorsal striatum. (as specified by Mawlawi *et al.*, 2001.)

input to the dorsal striatum). The putamen curves around the globus pallidus (which contains both internal or medial and external or lateral portions), and these two structures have together been referred to as the "lentiform" (Latin: "lentil-shaped") nucleus. Historically, the boundaries of the NAcc have not been clearly delineated anatomically (see, for example, the Talairach atlas (Talairach and Tournoux, 1988)), leading researchers to refer to the same region as both the NAcc and the head of the caudate. Finally, many researchers simply divide the striatum into lower (ventral) versus higher (dorsal) sections, with the ventral striatum encompassing the NAcc and lower portions of the putamen and caudate (and olfactory tubercle), while the dorsal striatum encompasses higher parts of the putamen and caudate (Gerfen and Wilson, 1996) (Figure 25.3). Researchers have recently proposed rotating this ventral–dorsal scheme by 45 degrees, forming a ventromedial to dorsolateral gradient (Voorn *et al.*, 2004).

Given these ambiguities, the anatomical definition of striatal components poses a challenge for neuroimagers. In humans, the caudate and putamen are clearly bounded on most sides by white matter or ventricles, but both share gray matter (and thus ambiguous) borders with the NAcc. Inspired by studies of neural connectivity, anatomical schemes for distinguishing ventral from dorsal striatum have been developed for positron emission tomography (PET) imaging (Drevets *et al.*, 2001; Mawlawi *et al.*, 2001) (Figure 25.3). Functional magnetic resonance imaging (fMRI) can provide finer spatial resolution, and thus the basis for identifying smaller striatal components delineated in comparative studies on the basis of cell type. According to one scheme, after orienting a brain in Talairach space, the NAcc can be distinguished from caudate dorsally by drawing a line from the tip of the internal capsule medially to the ventricle, and from the putamen laterally by drawing a line from the tip of the internal capsule ventrally to white matter (Breiter *et al.*, 1997). The anterior limit suggested by these authors extends to white matter, while the posterior limit was not specified, but has been set at or anterior to the anterior commissure in other studies (Haber and McFarland, 1999; Knutson *et al.*, 2001a; Figure 25.3).

Distinguishing between different striatal components is important, because these components connect to different brain regions and thus may support distinct functions. Early lesion studies suggested that different parts of the striatum projected to different parts of the prefrontal cortex via thalamic relays, which then projected back to the striatum, forming "loops" (Alexander *et al.*, 1986). Primate anatomical tracing studies have supported this scheme, demonstrating that the NAcc is connected to medial and ventral parts of the prefrontal cortex, while medial regions of the caudate and putamen connect to more dorsal regions of the anterior cingulate, and dorsal regions of the caudate and putamen connect to the dorsolateral prefrontal cortex (Haber, 2003). In addition, these studies indicated some overlap in projections, such that ventral loops connected to more dorsal loops in an ascending spiral running from ventromedial (i.e., NAcc/orbitofrontal cortex) to more dorsolateral (i.e., caudate tail/motor cortex) regions (Figure 25.4). Recent diffusion tensor imaging studies have verified this ascending spiral scheme of striatal-prefrontal connectivity in humans (Lehericy *et al.*, 2004). Thus, the striatum indirectly projects to the prefrontal cortex via inhibitory (i.e., GABAergic) projections that run through the subthalamic nucleus, pallidum, and thalamus. Reciprocally, a large part of striatal input comes from the frontal cortex via excitatory (i.e., glutamatergic) projections. Additionally, subcortical regions including the anterior insula, basolateral amygdala,

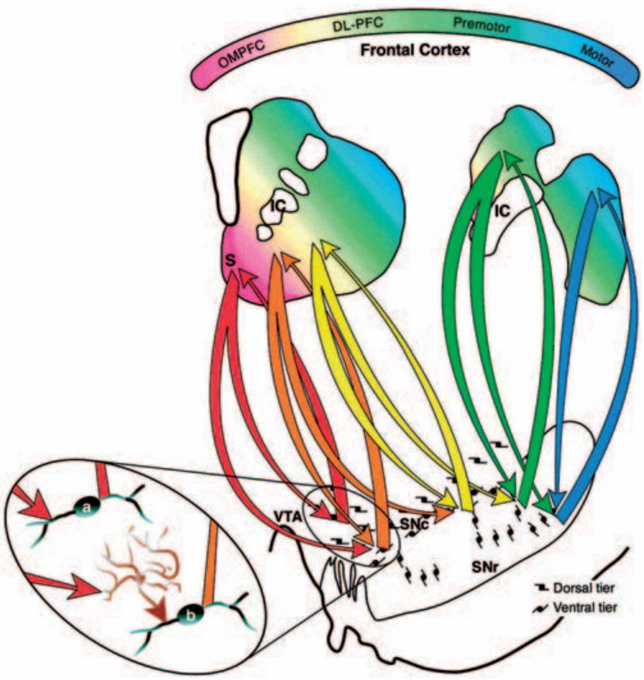


FIGURE 25.4 Ascending spirals of connectivity of the striatum to midbrain (downwards) and prefrontal cortex (upwards) (Reproduced from Haber *et al.* (2000), with permission. VTA; Ventral tegmental area, S; shell of the nucleus accumbens, SNc/SNr; Substantia Nigra compacta/reticulata, IC; internal capsule, OMPFC; orbitomedial perfrontal cortex, DLPFC: dorsolateral prefrontal cortex).

and hippocampus project to the ventral striatum (specifically, the shell of the NAcc) via excitatory (i.e., glutamatergic) projections (Friedman *et al.*, 2002). Finally, the ventral striatum sends prominent inhibitory (i.e., GABAergic) projections to the hypothalamus (Groenewegen *et al.*, 1999).

In addition to input and output involving simple excitatory (e.g., glutamate) and inhibitory (e.g., GABA) amino acid neurotransmitters, which act at specific synapses on a sub-second timescale, striatal activity is also modulated by larger biogenic amine and peptidergic neurotransmitters, which are likely to operate more broadly over multiple synapses on slower timescales of seconds or more. Of the biogenic amines, two midbrain dopaminergic nuclei prominently project to the striatum. Specifically, the more ventral and medial ventral tegmental area (VTA) nucleus projects to the ventral striatum, while the more dorsal and lateral substantia nigra (SN) nucleus projects to the dorsal striatum. Echoing the ascending spiral theme of striatal–prefrontal connectivity described above, the striatum sends descending inhibitory (GABAergic) projections to midbrain dopaminergic nuclei in a medial to lateral spiraling scheme (Haber *et al.*, 2000), implying that the striatum can exert some control over dopamine firing (Figure 25.4). Noradrenergic projections from the locus coeruleus of the midbrain to the striatum are much sparser, and primarily target the caudal NAcc shell and medial caudate adjacent to the stria terminalis, but not other striatal components (Berridge *et al.*, 1997). The caudal NAcc shell also receives the densest serotonergic projections from the dorsal raphe nucleus of the brainstem, with few serotonergic projections to the dorsal caudate, and sparse but evenly distributed projections to the putamen (Jacobs and Azmitia, 1992). Finally, tonically active neurons reside in and release acetylcholine in the striatum. Many peptides also innervate the striatum, mostly emanating from subcortical midline projections. While space limits preclude a full review of these peptidergic projections, comprehensive summaries can be found elsewhere (Nieuwenhuys, 1985; Holt *et al.*, 1997).

In summary, the human striatum connects to the rest of the brain (and particularly the prefrontal cortex) in a manner consistent with the notion that different components subserve different functions. Specifically, ventral striatal regions (i.e., NAcc, ventral caudate, and medial putamen) reciprocally target ventromedial cortical and subcortical regions implicated in emotion and motivation, while more dorsolateral striatal components (i.e., dorsolateral caudate and putamen) target dorsolateral cortical and subcortical regions implicated in movement and memory. The ascending loop trajectory implies that information flows through the striatum in a ventral to dorsal direction. This connectivity also implies that the striatum is ideally situated to coordinate valuation and subsequent action (Mogenson *et al.*, 1980).

FUNCTION

As early as 1664, British physician and scholar Thomas Willis observed striatal lesions in patients and speculated about function, noting "For here, … the animal Spirits, preparing for the force of the thing willed, are directed into appropriate nerves" (reprinted in Finger, 1994). Lesion studies still provide one of the few authoritative means of determining whether a structure provides a necessary (but not sufficient) substrate for a given behavioral function. A comprehensive meta-analysis of consequences of striatal lesions indicated that patients with lesions of the putamen (and globus pallidus) primarily presented with motor disturbances (e.g., dystonia), while patients with lesions of the caudate primarily presented with "behavioral" disturbances related to low motivation (e.g., abulia) (Bhatia and Marsden, 1994).

Since NAcc was not distinguished from caudate in this review, some of the caudate lesions may also have included NAcc lesions. Overall, consistent with Willis' early speculation, ventral striatal lesions were more likely to compromise motivation, while dorsal striatal lesions were more likely to compromise motor behavior.

A second historic body of lesion findings arises from characterization of diseases that compromise dopaminergic input to the striatum (rather than lesions of the striatum itself), such as Parkinson's disease. Typically, Parkinsonian degeneration progresses from the substantia nigra to the ventral tegmental area, and thus depletes dopamine from the dorsal striatum first and the ventral striatum second. Dorsal striatal dopamine depletion compromises affected individuals' ability to initiate or control movement. Patients also show cognitive deficits, however, particularly in the domain of feedback-dependent learning (Packard and Knowlton, 2002). Thus, in addition to examining motor deficits, researchers of these disorders have begun to explore motivational variables. Based on these findings and animal lesion studies, one influential theory describes the ventral striatum as primarily converting motivation to motor behavior (Mogenson *et al*., 1980).

Beyond investigating the long-term impact of lesions, scientists can directly record the rapid firing of neurons. As with lesion studies, researchers have typically focused either on the firing of neurons in the striatum (primarily GABAergic medium spiny output neurons), or on dopaminergic projections to these striatal neurons. Seminal primate studies have established that dopaminergic projections to the striatum fire in response not only to unexpected reward delivery, but also to reward cues, and decrease firing when predicted rewards do not occur – all consistent with a role in reward prediction (Schultz *et al*., 1997; see also Chapters 21 and 22 of this volume). However, these findings specify neither which striatal regions the dopamine neurons target, nor whether dopamine release varies across these targets. Primate studies of striatal neural activity suggest regional variation, with ventral striatal sensitivity to anticipated reward magnitude (Cromwell and Schultz, 2003) and dorsal striatal sensitivity to dynamically changing reward outcomes in the caudate (Hikosaka *et al*., 1989) and to habitual movements in the putamen (Rolls, 1994). While these correlational firing patterns imply that different striatal components might subserve different functions, causal manipulations of the activity of single neurons are both difficult to implement and often have little effect on overt behavior. Since other chapters prominently cover primate electrophysiology studies, we focus below primarily on studies of rodents and humans.

While lesions last forever and neural firing takes milliseconds, choices are made on an intermediate timescale of seconds. Over the past decade, technological advances now enable researchers to measure subcortical dopamine release (or correlated changes in activity) on a second-to-second timescale. Some methods provide greater chemical sensitivity but reduced temporal resolution, while other more recent methods provide greater temporal resolution but reduced chemical sensitivity. Remarkably, and possibly due to improvements in temporal resolution, findings in rats and humans have begun to converge.

Microdialysis

Microdialysis allows researchers to sample neurochemical activity in behaving animals (Zetterstrom *et al*., 1981). In microdialysis, a probe consisting of two concentric tubes with a membrane at the end is placed in a brain region. Fluid passes from the outer tube to the inner tube, and molecules in the extracellular space that are small enough to pass through the membrane are sampled. Microdialysis has increased neurochemical specificity but reduced spatial and temporal resolution relative to voltammetric methods (discussed below). For instance, over the past decades, researchers have sampled neurochemicals in regions as small as 0.3 millimeters every 2 minutes (Westerink, 1995; Salamone, 1996). The recent advent of capillary- and microfluidic-based systems will allow the collection of smaller samples, which may significantly improve temporal resolution to a scale of just a few seconds for amino acid and peptidergic neurotransmitters (Kennedy *et al*., 2002). Nonetheless, because microdialysis probes are larger than electrochemical probes, they can cause greater tissue displacement.

Several microdialysis studies indicate that feeding and drinking significantly increase NAcc dopamine release (Hernandez and Hoebel, 1988; Radakishun *et al*., 1988; Yoshida *et al*., 1992; Young *et al*., 1992; Di Chiara *et al*., 1999), with a few exceptions (Sabol *et al*., 1990; Cenci *et al*., 1992). Feeding and drinking also increase dopamine release in dorsal striatal regions such as the caudate (Young *et al*., 1992), and other mesolimbic regions including the mesial prefrontal cortex (MPFC) (Cenci *et al*., 1992; Young *et al*., 1992) and the VTA (Yoshida *et al*., 1992). While all of these studies included food- or water-restricted rats, even in non-restricted rats, consumption of a palatable food can increase NAcc dopamine release beyond that evoked by consumption of a non-palatable food

(Martel and Fantino, 1996). Given the limited temporal resolution of microdialysis, these studies cannot distinguish anticipation of food or liquid rewards from consumption – a shortcoming which can be partially addressed with experimental manipulations. For instance, presentation of food pellets every 45 s (which also generates anticipatory behavior) increases NAcc dopamine release more than massed presentation of the same amount of food (McCullough and Salamone, 1992). Further, observation of food prior to eating elicits dopamine release in the NAcc (and MPFC), prior to dopamine release elicited by eating, and food-specific satiety blunts this anticipatory release (Ahn and Phillips, 1999). NAcc dopamine release also robustly increases during mating in both males and females (Pfaus *et al*., 1990), as well as in experienced males when presented with an estrous female behind a screen (Damsma *et al*., 1992). Compared with other motorically engaging activities (Martin and Myers, 1976), NAcc dopamine release increases more during copulation than in response to a novel environment or vigorous locomotor activity (Damsma *et al*., 1992).

Striatal dopamine release is not limited to rewarding circumstances, since many researchers also report increases in NAcc dopamine release during presentation of punishing stimuli such as foot shock (Sorg and Kalivas, 1991), tail shock (Abercrombie *et al*., 1989), or restraint stress (Imperato *et al*., 1992). Cues that predict footshock can also increase NAcc dopamine release (Young *et al*., 1993). While most of these studies did not compare relative levels of dopamine release in response to rewarding versus punishing stimuli, they do suggest that presentation of punishing stimuli can elicit dopamine release. However, some of these studies suggest that NAcc dopamine release is greater during the offset than during the onset of the aversive stimuli, which may still support an appetitive role for NAcc dopamine, even under aversive circumstances. For example, greater increases in NAcc dopamine release are often observed when animals work to actively avoid footshock than when they actually receive footshock (McCullough *et al*., 1993). Further, while NAcc dopamine increases during an initial session of restraint stress (40 min), this effect diminishes over the course of several repeated sessions, but increases after the offset of restraint stress, and these effects grow more pronounced over several repetitions (Imperato *et al*., 1992). Finally, some aversive cues reduce rather than increase NAcc dopamine, such as taste cues associated with the induction of nausea (Mark *et al*., 1991).

In a "unified" interpretation of NAcc dopamine release, Ikemoto and Panksepp (1999) proposed that, under aversive circumstances, NAcc dopamine increases as a function of an animal's perception that it can escape a stressor, which presents a positive rather than negative incentive. According to this theory, the anticipatory nature of NAcc dopamine release may especially hold for the NAcc shell, which is associated with incentive processing more than the NAcc core, which is associated with generating motor output. This "escapability" hypothesis is consistent with the findings of a study in which novel appetitive but not aversive stimuli increased dopamine release in the NAcc shell (Bassareo *et al*., 2002), while both novel appetitive and aversive stimuli increased dopamine in the NAcc core and MPFC. Further, when aversive stimuli were removed, dopamine increased 10 minutes later, but these increases were blocked by reintroduction of the aversive stimulus. A second piece of support for the escapability hypothesis comes from a study in which administration of footshock increased dopamine in the MPFC but not in the NAcc. NAcc dopamine only increased over repeated testing sessions, as rats learned to avoid the shock (Wilkinson *et al*., 1998).

In summary, although microdialysis lacks the temporal resolution to reveal phasic changes in dopamine release, experiments designed to distinguish appetitive (or anticipatory) from consummatory (or outcome) phases of reward processing indicate increased NAcc dopamine release during reward anticipation. Interestingly, dopamine release increases in response to the offset of aversive stimulation in the NAcc shell (but not other regions like the MPFC).

Voltammetry

During the last quarter of the twentieth century, researchers refined electrochemical methods for indexing dopamine activity (i.e., chronoamperometry and cyclic voltammetry) (Kissinger *et al*., 1973). In these methods, researchers apply an electrical potential to an electrode to oxidize nearby molecules in the extracellular space. Molecular identity can then be inferred from "redox" properties of the oxidized material (i.e., a combination of peak oxidation, peak reduction potential, and electron transfer kinetics) (Phillips and Wightman, 2003). Relative to microdialysis, advantages of electrochemical methods include increased spatiotemporal resolution (i.e., 5- to 30-micrometer diameter carbon fibers that can typically sample at 50–100 ms) and decreased disruption of surrounding tissue (Peters *et al*., 2004). The primary disadvantage involves decreased chemical selectivity, particularly if different molecules with similar electroactive potentials are near the sensor (Wightman and Robinson, 2002).

Additionally, background pH changes associated with neural activity can interfere with dopamine detection in chronoamperometry (which produces "dopamine-like" signals), which can be corrected for in voltammetry (Wightman and Robinson, 2002). Thus, while microdialysis indexes tonic changes in dopamine, electrochemical methods index phasic changes in dopamine, which usually result from dopamine "overflow" outside the synapse into the extracellular space following burst-firing of dopamine neurons.

Cyclic voltammetry studies (1 sample/1 s) have revealed phasic increases in dopamine release in the NAcc shell (<10 s) as rats enter novel environments, and these phasic increases habituate upon subsequent reintroductions (Rebec *et al.*, 1997). Using chronoamperometry (1 sample/2 s), phasic increases in NAcc dopamine-like signals have been documented in food-restricted rats as they anticipated responding for milk droplets. These dopamine-like signals decreased after responding with a lever-press and during milk consumption (Richardson and Gratton, 1998). In a related chronoamperometry study conducted at a slower sampling rate (1 sample/60 s), NAcc dopamine-like signals increased after presentation of a cue that predicted a palatable liquid meal, but remained elevated during food consumption (Phillips *et al.*, 1993). Using pulsed voltammetry (1 sample/120 s), investigators noted an increase in dopamine in the NAcc shell and medial part of the dorsal striatum in response to presentation of a novel appetitive odor (banana) (Besson and Louilot, 1997; Jeanblanc *et al.*, 2002). This increase was prevented during a second testing session by pairing the smell with a nausea-inducing injection of lithium chloride, relative to the increase observed in a control group who had the smell paired with an injection of vehicle.

As with microdialysis, pulsed voltammetry (1 sample/60 s) has revealed increases in NAcc dopamine-like signal in male rats exposed to an estrus female, ovariectomized female, or another male rat (Mas *et al.*, 1995). However, a chronoamperometric study with greater temporal resolution (1 sample/5 s) revealed greater NAcc increases in dopamine-like signals in response to bedding from an estrus female, compared to bedding from an ovariectomized female, or another male rat (Mitchell and Gratton, 1991). Increases in dopamine-like signal were also greater in the NAcc than in dorsal striatal regions. Finally, cyclic voltammetry studies (1 sample/0.1 s) have now demonstrated robust increases in NAcc dopamine in male rats exposed to a receptive female rat (Robinson *et al.*, 2001), and when male rats were initially introduced to conspecifics. These phasic increases were often followed by increased social investigation (Robinson *et al.*, 2002). Other cyclic voltammetry studies have clearly demonstrated increases in NAcc dopamine prior to self-administration of food (Roitman *et al.*, 2004) and cocaine (Phillips *et al.*, 2003).

Taken together, the few voltammetric studies with second-to-second temporal resolution (e.g., <1 s) indicate not only that NAcc dopamine increases in anticipation of reward (particularly in the shell), but also that these increases are more robust than those that occur during consumption. The findings are consistent with rat lesion studies linking the ventral striatum to acquisition of appetitive behavior, and the dorsal striatum to the maintenance of habitual behavior associated with consuming rewards (Robbins *et al.*, 1989). The findings also cohere remarkably well with primate studies on the firing of dopamine neurons in response to rewarding incentives (Schultz *et al.*, 1997). Specifically, ventral striatal dopamine release occurs in response to unexpected reward cues, but tracks reward cues once their predictive association is learned (Figure 25.5). Fewer studies have focused on the dorsal striatum or on dopamine release in response to aversive stimuli. Such comparisons may provide promising lines of inquiry for future research.

Neurally, humans differ from rats. For instance, humans have a more highly elaborated prefrontal cortex. However, subcortically (and especially in the striatum), differences are less pronounced. Although

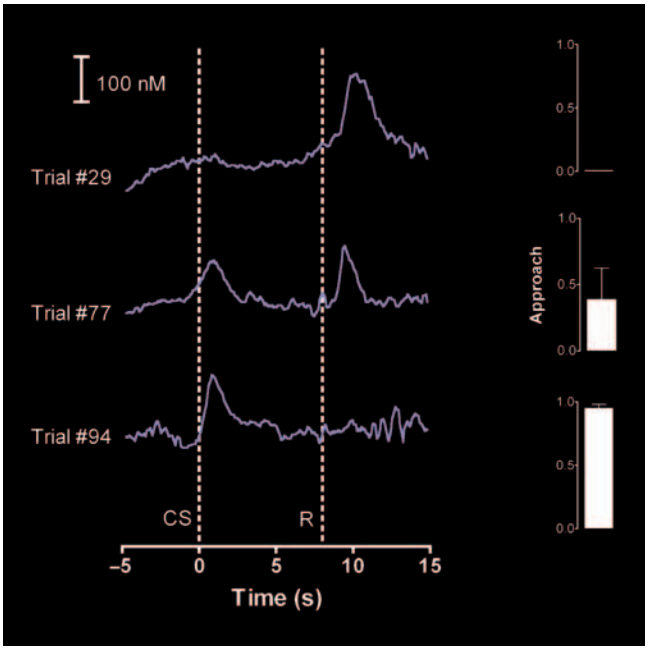


FIGURE 25.5 Nucleus accumbens dopamine release shifts from juice delivery to reward cue presentation over the course of training, and correlates with approach to the cue (CS, conditioned stimulus; R, reward). J.J. Clark and P.E.M. Phillips (unpublished data).

microdialysis and *in vivo* cyclic voltammetry of dopamine release have not yet been applied to humans, two parallel neuroimaging methods have been implemented over the past decade. Specifically, ligand-based PET scanning allows researchers to make inferences about dopamine release, but only on a timescale of ~60 minutes. On the other hand, FMRI scanning assesses changes in blood oxygenation rather than dopamine release, but on a timescale of seconds. Humans also differ psychologically from rats. Conveniently for neuroeconomics research, humans have devised a consensual symbolic representation of value called "money." Use of monetary incentives not only allows researchers to compare neural responses to scaled gains versus losses, but also facilitates analysis of economic choice (e.g., deciding to purchase a product or share in a company). Thus, here we primarily consider neuroimaging studies that focus on monetary incentives and related choices.

Positron Emission Tomography (PET)

Positron emission tomography (PET) includes a variety of techniques that enable researchers to visualize metabolic and neurochemical changes in brain activity. In metabolic PET, researchers typically inject radioactively tagged oxygen or glucose into subjects, which is taken up into active brain regions as subjects perform a task. The tagged positron decays, emitting two electrons at 180 degrees that can be coincidentally detected with a PET camera. The two most popular metabolic PET techniques measure changes in cerebral blood flow (rCBF) with ^{15}O-water or in regional cerebral metabolic rate for glucose utilization (rCMRglc) with ^{18}F-fluorodeoxyglugocse (FDG).

In ligand-based PET, researchers inject a radioactively tagged ligand (or chemical that mimics a neurotransmitter) that travels to the brain and binds to specific receptors. As subjects perform a task, they release neurotransmitter, which displaces the tagged ligand. Relative displacement can be detected when a task condition involving neurotransmitter release is contrasted with a control condition in which the neurotransmitter is not released. Thus, unlike FMRI, PET can facilitate neurochemical inference and does not suffer from artifacts in regions near tissue boundaries (see the next section). However, PET typically has less spatial (~8 mm) and temporal resolution (>2 min) than FMRI.

The effect of monetary incentives on neural activity was initially investigated with regional cerebral blood flow PET (Thut *et al.*, 1997). In this study, subjects performed a delayed go–no go task, in which correct answers were rewarded either with money or with a non-monetary affirmative prompt. Monetary reward increased activity in the midbrain, thalamus, OFC, and dorsolateral prefrontal cortex to a greater extent than did the prompt. The authors concluded that this pattern of activation might reflect the assessment of consequences in goal-directed behavior. In a second PET study using a pattern-recognition task, monetarily rewarded conditions activated the left cerebellum, midbrain, striatum, anterior cingulate, and MPFC, whereas symbolically rewarded conditions activated only the anterior cingulate and MPFC (Kunig *et al.*, 2000).

Given the strong association between reward and dopamine release in animal research, it is no surprise that dopamine ligand-based PET studies of humans have also focused on the striatum. Interpretation of results from these studies is limited by two factors, however. First, it is unclear whether reward processing increases dopamine release in non-striatal regions (e.g., amygdala, thalamus, and cortex) since the relatively low density of dopamine receptors in these areas precludes visualization with current variants of ligand-based PET. Second, the D2 agonist raclopride (the radiotracer used in most dopamine PET studies) is the only ligand that has been shown to be displaceable in real time by endogenous neurotransmitter in response to behavioral manipulations (although other more sensitive ligands are in development).

The first study to investigate the relationship between dopamine release and non-pharmacological reward utilized raclopride displacement while subjects played a goal-directed motor task (videogame) (Koepp *et al.*, 1998). As subjects gained points during the game, there was a decrease in raclopride binding in the ventral striatum (NAcc and caudate), indicating increased dopamine release and binding at D2 receptors. These results are similar to the increased raclopride displacement in the ventral striatum observed following the presentation of unexpected monetary gain (Pappata *et al.*, 2002). Although the amplitude of this effect was moderate, there was no detectable change in dopamine release following unexpected monetary loss, which represents a condition that should control for arousal. Further raclopride displacement studies suggest complex changes in striatal dopaminergic activity during a gambling task involving monetary reward (Zald *et al.*, 2004).

Thus, while metabolic PET studies have not consistently implicated the striatum in processing of monetary reward, ligand-based PET studies of DA displacement do implicate the striatum, and at least one study suggests that equivalent punishing monetary incentives do not have similarly robust effects. Overall, due to limited temporal resolution, researchers have found it difficult to distinguish reward anticipation from outcome using PET – an issue that has been addressed with event-related fMRI.

Functional Magnetic Resonance Imaging (fMRI)

fMRI enables researchers to visualize changes in blood oxygenation. In the blood oxygen level dependent (BOLD) effect (hereafter called "activation"), approximately 4–6 seconds after neural activity occurs, an excess of oxygenated blood is delivered to that brain region. This localized pooling of oxygenated hemoglobin creates a transient magnetic inhomogeneity that can be detected with a magnetic resonance scanner. fMRI activation correlates more closely with changes in postsynaptic membrane potential than with changes in presynaptic firing, and so has been postulated to index the combined input to a brain region (Logothetis *et al*., 2001). Unlike ligand-based PET imaging, fMRI cannot presently index specific neurochemical changes, although combined pharmacological and fMRI studies may eventually elucidate these links (Knutson and Gibbs, 2007). Additionally, many brain areas of central interest to reward researchers (e.g., the NAcc, orbital frontal cortex, and MPFC) lie near tissue boundaries (i.e., next to ventricles and sinuses), which can cause artifacts unless care is taken to minimize magnetic inhomogeneities (e.g., by use of special pulse sequences or acquisition parameters). Relative to PET, however, fMRI affords increased spatial resolution (e.g., as small as $1\,\text{mm}^3$ versus $8\,\text{mm}^3$) and, importantly, substantially increased temporal resolution (e.g., seconds rather than minutes). fMRI also has advantages of safety and convenience – since blood itself provides the signal, researchers need not inject radioactive compounds or other agents into subjects prior to scanning. These advantages have made FMRI one of the most popular methods for visualizing changes in the activation of small subcortical structures.

Shortly after its development, researchers began to use fMRI to examine neural responses to both primary (e.g., touch, juice, food, odors) and secondary (e.g., money) incentives. Primary and secondary incentives have yielded overlapping patterns of activation, but the extent of their similarity has not yet been fully characterized (see Chapter 24 of this volume; O'Doherty, 2004). Based on their direct relevance to neuroeconomic questions, we focus here on studies that utilized monetary incentives.

Adopting a successful strategy from fMRI studies of vision research, initial fMRI studies attempted to "localize" or correlate regionally specific patterns of brain activation with responses to monetary incentives in the absence of choice. In an initial fMRI study utilizing a cued reaction-time task, trials involving monetary gain or loss activated striatal regions (including caudate and putamen) as well as other mesolimbic and motor regions (i.e., insula, MPFC, supplementary motor area, and motor cortex), relative to identical trials that lacked monetary incentives (Knutson *et al*., 2000). Trials involving monetary loss additionally activated anterior thalamus and anterior cingulate. However, incentive anticipation and outcomes were not separately modeled in this experiment, and the investigators were not able to acquire images in the anterior portion of the prefrontal cortex. In a second study utilizing a gambling task, trials involving financial gains activated striatal regions in the putamen (as well as the midbrain), while trials involving financial losses activated other subcortical regions (in the medial temporal lobe) (Elliott *et al*., 2000). Although this study identified distinct patterns of activation for gambles involving gains versus losses, the study's design did not allow separate analysis of incentive anticipation and outcomes. In a third study using a different gambling task that controlled for anticipation, gain outcomes elicited sustained activation of striatal regions in the caudate relative to loss and neutral outcomes (Delgado *et al*., 2000). In a fourth study using a reversal learning task, gain versus neutral outcomes activated medial prefrontal regions, while loss versus neutral outcomes activated more lateral prefrontal regions (O'Doherty *et al*., 2001). However, in this study investigators acquired data from the prefrontal cortex only, but not the striatum.

Subsequent fMRI studies utilizing monetary incentives attempted to dissociate incentive anticipation from outcomes. One study using a gambling task revealed activations of NAcc and other regions (medial OFC and cerebellum) for gain versus neutral gambles, but activations of other regions (temporal lobe, lateral OFC, and cuneus) for the loss versus neutral gambles (Breiter *et al*., 2001). The authors concluded that both anticipation and outcomes related to monetary gains activated ventral striatum as well as other mesolimbic regions (medial OFC, and hypothalamus). A second study using a cued reaction time task revealed proportional activation in the NAcc, medial caudate, and anterior thalamus during anticipation of increasing monetary gains, but only proportional activation in the medial caudate and anterior thalamus during anticipation of increasing monetary losses (Knutson *et al.*, 2001a). NAcc activation also correlated with individual differences in self-reported positive arousal in response to seeing large ($5.00) reward cues. Based on these findings, Knutson *et al.* (2001a) proposed that NAcc activation specifically and proportionally correlated with anticipated gain (Figure 25.6). Follow-up studies utilizing a similar design replicated the association between NAcc activation and gain anticipation, but further indicated that gain versus non-gain outcomes instead recruited medial caudate, MPFC, and posterior cingulate activation (Knutson *et al.*, 2001b, 2003).

These early findings foreshadowed a steep linear increase in the number of fMRI publications utilizing monetary incentives, continuing to the present (Knutson and Cooper, 2005; Delgado, 2007). Some studies have focused on anticipation, others on outcome, and still others on learning. With respect to anticipation, researchers began to examine whether ventral striatal activation correlated with other aspects of anticipated gains, such as probability and delay. For instance, one study examined the foundational economic construct of expected value by independently manipulating the magnitude and probability of cued gains and losses. In a cued reaction-time task, anticipated gain magnitude times probability correlated with activation in midbrain, ventral striatum, and MPFC (Figure 25.7). However, decomposition of magnitude and probability effects suggested that while ventral striatal activation was sensitive primarily to expected gain magnitude, MPFC was additionally sensitive to expected gain probability (Knutson *et al.*, 2005). Studies using gambling tasks similarly indicated that ventral striatal activation during gain anticipation increased as a function of magnitude, but peaked at an intermediate probability value (Dreher *et al.*, 2006; Preuschoff *et al.*, 2006 Cooper and Knutson, 2008). Other studies using gambling tasks, however, suggest that ventral striatal activation correlates with both the magnitude and the probability of anticipated gains (Abler *et al.*, 2006; Yacubian *et al.*, 2006). While all of these findings indicated that ventral striatal activation correlates with the magnitude of anticipated gain, the reasons for discrepant findings with respect to probability remain unclear. According to one possibility, a phasic signal may pass through the ventral striatum that correlates with both magnitude and probability, followed by a tonic signal that primarily correlates with magnitude (see also Chapter 21 of this volume). Alternatively, it is possible that a

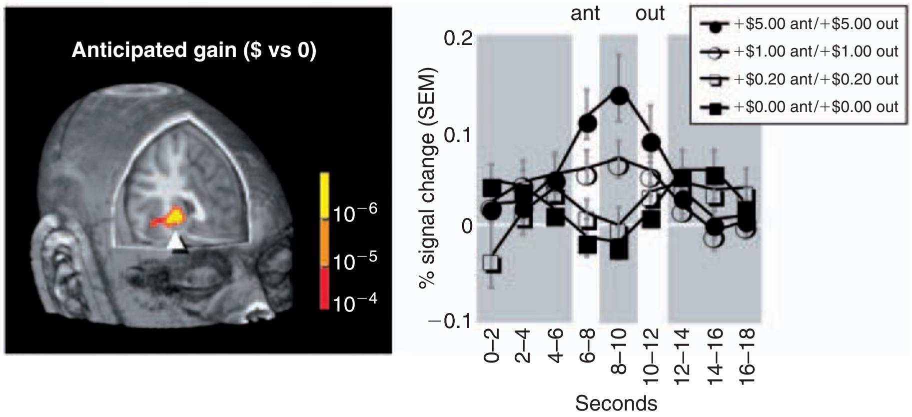


FIGURE 25.6 Anticipation of monetary gains of increasing magnitude elicits proportionally increasing NAcc activation. (adapted from Knutson *et al.*, 2003.)

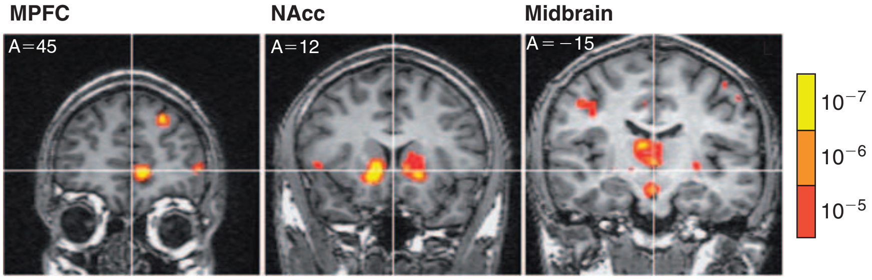


FIGURE 25.7 Neural activation correlated with expected value (anticipated monetary gain magnitude times probability. Reproduced from Knutson *et al.*, (2005), with permission.

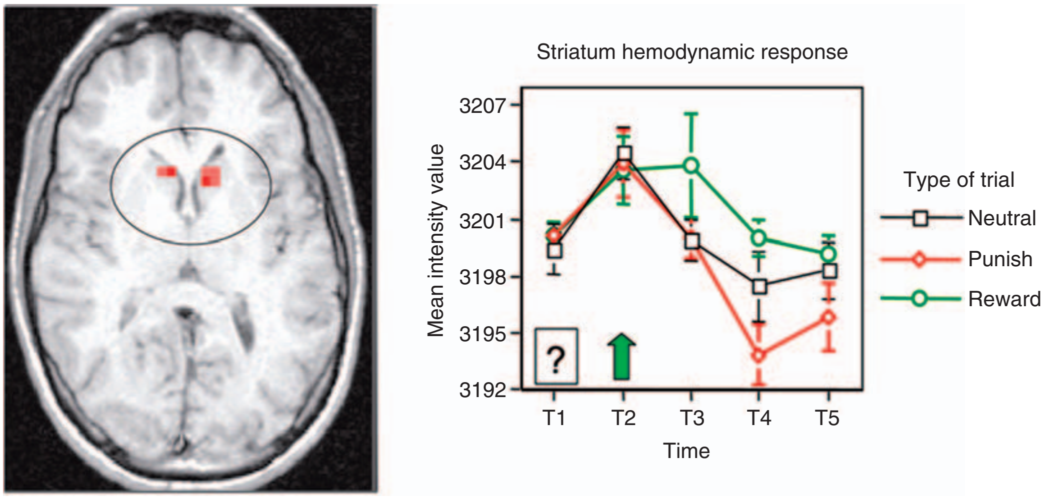


FIGURE 25.8 Caudate response to incentive outcomes. Reproduced from Delgado *et al*. (2000), with permission.

signal from the MPFC reduces ventral striatal activation in response to non-gain outcomes, which are more prevalent in low-probability trials, and that some of this reduction is being modeled during anticipation periods. Future research will doubtless resolve these issues. Other studies have focused on the delay of anticipated gains, but these designs have not yet distinguished between anticipation, choice, and outcome phases of incentive processing. In one delayed discounting task, ventral striatal, MPFC, and posterior cingulate activation was correlated with consideration of choices which featured an immediate and delayed option versus two delayed options (McClure *et al*., 2004). However, in a second study using a delayed discounting task that held the immediate option constant, activation in these same regions correlated with the discounted value of the delayed option (i.e., a combination of magnitude and delay; Kable and Glimcher, 2007).

With respect to incentive outcomes (i.e., when the probability of a gain or loss collapses from an intermediate value to either 1 or 0), researchers began to examine whether caudate activation correlated with outcome value, and whether this response to outcomes had implications for future choice. Further studies replicated and extended initial work on caudate responses to outcomes in which subjects received or did not receive monetary gains, indicating a parametric response of caudate activation to outcome valence and magnitude (Delgado *et al*., 2000, 2003, 2004) (Figure 25.8). Emerging evidence further indicated that caudate activation correlates with relative as well as absolute monetary gain outcomes (i.e., counterfactual or "fictive" outcomes in which an obtained outcome is compared to an obtained alternative outcome) (Kuhnen and Knutson, 2005; Nieuwenhuis *et al*., 2005; Lohrenz *et al*., 2007). This research has begun to forge connections with a separate literature that has historically focused on cognitive rather than incentive feedback (Elliott *et al*., 1997; Poldrack *et al*., 1999), indicating overlapping patterns of activation of the caudate, but with more robust recruitment in response to incentive feedback (Tricomi *et al*., 2006).

While these incentive outcome findings implicated the caudate in representing the value of incentive outcomes, they did not clarify whether increased caudate activation resulted from gain outcomes themselves or from the updating of action associations that produced gains (more akin to a traditional "reinforcement" signal). Some evidence implicated increased caudate activation when outcomes informed future choices in active but not passive tasks (O'Doherty *et al*., 2004), as well as when outcomes informed future behavior (Tricomi *et al*., 2004). These findings are consistent with rodent studies implicating the caudate in the representation of action values (Yin *et al*., 2005), but the putamen in the representation of more habitual actions (Yin *et al*., 2006). The action valuation account is also consistent with findings from recent studies in which caudate responses diminished as reward contingencies became predictable, and thus carried less information about the next best action (Haruno *et al*., 2004; Delgado *et al*., 2005a).

In summary, fMRI studies of monetary incentive processing consistently implicate the ventral striatum (including the NAcc) in the processing of monetary gain but not loss. A number of studies are beginning to suggest that activation of the NAcc occurs most robustly during anticipation of gains, while activation of the caudate occurs most robustly in response to gain outcomes that call for future action, consistent with the ascending spiral of connectivity implied by structural studies (Haber *et al.*, 2000).

PREDICTION

If striatal activation correlates with the representation of subjective value, then it should also contribute to subsequent choice (both economic and social) and perhaps even to memory. Indeed, direct infusion of dopamine-releasing agents into the ventral (but not dorsal) striatum of rats elicits increased approach to stimuli that have previously predicted reward (Parkinson *et al.*, 1999), and unconditionally evokes appetitive (but not aversive) behavior (e.g., forward locomotion, sniffing, and 50-kHz ultrasonic vocalizations) (Burgdorf *et al.*, 2001). Conversely, microstimulation of the caudate during outcome processing (but not during reward anticipation) improves monkeys' ability to learn the next appropriate response for rewards (Nakamura and Hikosaka, 2006). Event-related fMRI studies in humans are now beginning to test whether striatal signals presage changes in choice and memory.

Studies have begun to suggest that activation of midbrain and ventral striatum can facilitate memory formation. In one study, cues that signaled monetary gain were better remembered than cues that did not 2 days after scanning, and this effect was correlated with cue-elicited activation of the midbrain, ventral striatum, and hippocampus (Wittmann *et al.*, 2005). In a second study, subjects better remembered neutral stimuli that followed cues signaling monetary gain for successful memorization, and this effect was mediated by cue-elicited activation of the midbrain, ventral striatum, and hippocampus (Adcock *et al.*, 2006). This growing body of research suggests that ventral striatal activation may influence memory formation, possibly by recruiting midbrain dopaminergic projections to hippocampal regions.

Other studies suggest that activation in these regions might contribute to economic choice. In one study, subjects chose between high- versus low-risk investments. Ventral striatal activation immediately prior to choice predicted that subjects would be more likely to switch to high-risk investments above and beyond informational variables (e.g., wealth, prior outcome, uncertainty) even when subjects' choices violated those of a rational actor (i.e., a Bayesian updating, risk-neutral agent) (Kuhnen and Knutson, 2005). In another study, subjects decided whether or not to purchase products at discounted prices. Ventral striatal activation not only correlated with preference while viewing products, but also predicted that subjects would be more likely to choose to purchase the product above and beyond self-reported preference (Knutson *et al.*, 2007). These studies suggested that naturally occurring (or endogenous) changes in ventral striatal activation precede and can be used to predict subsequent choice. A third study further suggested that this influence could be externally (or exogenously) controlled. In this study, heterosexual male subjects saw positive (i.e., erotic), negative (i.e., snakes and spiders), or neutral (i.e., office supplies) pictures prior to choosing between high- versus low-risk gambles of equal expected value. Presentation of positive pictures increased switching to the high-risk gamble, and this effect was partially statistically mediated by increases in ventral striatal activation (Knutson *et al.*, 2008) (Figure 25.9). Together, these studies are consistent with a causal account in which physiological events indexed by ventral striatal activation influence subsequent economic decisions.

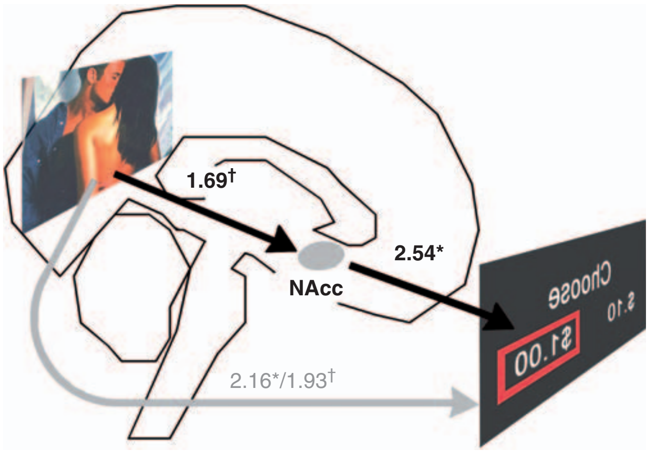


FIGURE 25.9 Activation of nucleus accumbens statistically mediates the influence of positive pictures on switching to a high risk financial gamble (Z-scores, $^{*}P < .025$, $^{\dagger}P < .05$, one–tailed). Knutson *et al*. (2008).

A third group of studies has implicated striatal activation in social decisions (see also Chapter 15 and 22 in this volume). For instance, ventromedial caudate activation not only occurred in response to reciprocated trust in a prisoner's dilemma game (Rilling *et al.*, 2002), but also predicted subjects' tendency to invest in a partner who had cooperated with them in the past in a trust game (King-Casas *et al.*, 2005). Thus, ventral striatal activation can come to represent future social gain. Consistent with this notion, information about reputation of social partners modulated the responsiveness of these regions to their behavior in the absence of any previous experience (Delgado *et al.*, 2005b).

IMPLICATIONS

In summary, consistent with the "ascending spiral" structural connections of the striatum, different parts of the striatum appear to subserve different functions. Further, emerging evidence suggests that striatal activation can be used to predict choice (and memory) within trials. These advances have resulted from improvements in the spatial and especially temporal resolution of neuroimaging techniques. Together, these findings reveal a picture of incentive processing in which the ventral striatum assesses expected gain and the dorsal striatum uses that estimate to inform future actions and cognitions.

Over the course of approximately a decade of research, we have learned lessons that might facilitate future investigations. First, valuation appears to be a dynamic and componential process. Second, different stages of this process recruit different striatal components, perhaps in an ascending progression (see also Chapter 24 of this volume). Third, spatiotemporal resolution is key for elucidating this process, and so further technical advances should facilitate theoretical developments. The findings have also forced a re-examination of some theoretical assumptions. First, neural processing of gains does not appear to be the opposite of processing of losses. Different circuits may handle these distinct incentive processing demands, enabling animals simultaneously to process potential gain and potential loss. Even within the striatum, the evidence for responsiveness to losses is weaker in the ventral striatum than in the dorsal striatum. Additional findings from our labs and others' suggest that other brain regions (the insula, for instance) may play a more prominent role in responses to loss. Second, the brain responds differently during anticipation of incentives than in response to incentive outcomes. While the ventral striatum appears more active during anticipation, the dorsal striatum appears to respond more robustly to outcomes, particularly if the outcomes inform what an animal needs to do next. Hopefully, re-examination of these assumptions will generate more robust assumptions that can support more inclusive and predictive theories of choice.

Mostly, we have learned how much we have yet to learn. For instance, other types of valuation beg for investigation. These include both positive incentives (i.e., probability, immediacy, certainty) and negative incentives (i.e., effort, risk, delay). Are these incentives processed by the same or by different neural circuits? If the latter, how does the brain integrate these components to inform future choices and expectations (i.e., learning)? Can researchers use these neural signals to better predict choice (e.g., beyond existing economic and psychological models)? How are the signals modulated by contextual economic and social variables? And how can people learn to better control these signals to promote choices most consistent with their long-term goals?

As implied throughout this review, if the past provides a guide to the future, technological innovations should continue to drive theoretical advances. Specifically, improvements in spatial and especially temporal resolution should improve inference about the function of different striatal components. Enhanced resolution will increase researchers' ability to compare neural signals to existing economic and psychological models, and to use these signals to predict economic choice and psychopathological symptoms. Enhanced spatiotemporal resolution will also facilitate bridge-building to animal research, as well as promoting integration of information from other measurement modalities (e.g., electrical and chemical). Eventually, the process approach may supplant the correlational approach.

A deep and revolutionary assumption underlies these predictions – choice cannot be understood without an understanding of subjective valuation. Past traditions in economics (e.g., revealed preferences) and psychology (e.g., behaviorism) assumed the opposite – that choice cannot be understood without translating (internal) subjective valuation to (external) objective stimuli (see Chapter 9 of this volume). Technology no longer forces us to adhere to this assumption. Subjective value stands at the nexus of sensory input and motor output, and tools that facilitate its measurement lie within our reach.

In conclusion, neuroscience findings suggest that valuation is a dynamic, componential, and ultimately subjective process. Technological advances now allow scientists to track these components of valuation, providing reason for hope rather than despair. Whether or not it is the site at which "animal spirits" transform sensory impressions into motor command, the striatum seems like a promising place to begin to search for the neural basis of subjective value.

Acknowledgments

During chapter preparation, BK was supported by National Institute of Health Grant AG030778 and a FINRA Investor Education Award, MRD was supported by National Institute of Health grant DA022998, and PEMP was supported by National Institutes of Health grants AG30775, DA21793, DA24140 and MH79292, and a NARSAD Young

Investigator Award. The authors thank Paul Glimcher and Jeffrey C. Cooper for helpful comments on earlier versions of the manuscript.

References

Abercrombie, E.D., Keefe, K.A., DiFrischia, D.S., and Zigmond, M.J. (1989). Differential effect of stress on in vivo dopamine release in striatum, nucleus accumbens, and medial frontal cortex. *J. Neurochem.* 52, 1655–1658.

Abler, B., Walter, H., Erk, S. *et al.* (2006). Prediction error as a linear function of reward probability is coded in the human nucleus accumbens. *NeuroImage* 31, 790–795.

Adcock, R.A., Thangavel, A., Whitfield-Gabrieli, S. *et al.* (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron* 50, 507–517.

Ahn, S. and Phillips, A.G. (1999). Dopaminergic correlates of sensory-specific satiety in the medial prefrontal cortex and nucleus accumbens of the rat. *J. Neurosci.* 19, RC29.

Alexander, G.E., DeLong, M.R., and Strick, P.L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* 9, 357–381.

Bassareo, V., De Luca, M.A., and Di Chiara, G. (2002). Differential expression of motivational stimulus properties by dopamine in nucleus accumbens shell versus core and prefrontal cortex. *J. Neurosci.* 22, 4709–4719.

Berridge, C.W., Stratford, T.L., Foote, S.L., and Kelley, A.E. (1997). Distribution of dopamine-beta-hydroxylase-like immunoreactive fibers within the shell subregion of the nucleus accumbens. *Synapse* 27, 230–241.

Besson, C. and Louilot, A. (1997). Striatal dopaminergic changes depend on the attractive or aversive value of stimulus. *NeuroReport* 8, 3523–3526.

Bhatia, K.P. and Marsden, C.D. (1994). The behavioural and motor consequences of focal lesions of the basal ganglia in man. *Brain* 117, 859–876.

Breiter, H.C., Gollub, R.L., Weisskoff, R.M. *et al.* (1997). Acute effects of cocaine on human brain activity and emotion. *Neuron* 19, 591–611.

Breiter, H.C., Aharon, I., Kahneman, D. *et al.* (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30, 619–639.

Burgdorf, J., Knutson, B., Panksepp, J., and Ikemoto, S. (2001). Nucleus accumbens amphetamine microinjections unconditionally elicit 50-kHz ultrasonic vocalizations in rats. *Behav. Neurosci.* 115, 940–944.

Cenci, M.A., Kalen, P., Mandel, R.J., and Bjorklund, A. (1992). Regional differences in the regulation of dopamine and noradrenaline release in medial frontal cortex, nucleus accumbens and caudate-putamen: a microdialysis study in the rat. *Brain Res.* 581, 217–228.

Cooper, J.C. and Knutson, B. (2008). Valence and salience contribute to nucleus accumbens activation. *NeuroImage* 39, 538–547.

Cromwell, H.C. and Schultz, W. (2003). Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J. Neurophysiol.* 89, 2823–2838.

Damsma, G., Pfaus, J.G., Wenkstern, D. *et al.* (1992). Sexual behavior increases dopamine transmission in the nucleus accumbens and striatum of male rats: comparison with novelty and locomotion. *Behav. Neurosci.* 106, 181–191.

Dawkins, R. (1989). *The Selfish Gene*. Oxford: Oxford University Press.

Delgado, M.R. (2007). Reward-related responses in the human striatum. *Ann. NY Acad. Sci.* 1104, 70–88.

Delgado, M.R., Nystrom, L.E., Fissell, C. *et al.* (2000). Tracking the hemodynamic response to reward and punishment in the striatum. *J. Neurophysiol.* 84, 3072–3077.

Delgado, M.R., Locke, H.M., Stenger, V.A., and Fiez, J.A. (2003). Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cogn Affect. Behav. Neurosci.* 3, 27–38.

Delgado, M.R., Stenger, V.A., and Fiez, J.A. (2004). Motivation-dependent responses in the human caudate nucleus. *Cerebral Cortex* 14, 1022–1030.

Delgado, M.R., Miller, M.M., Inati, S., and Phelps, E.A. (2005a). An FMRI study of reward-related probability learning. *NeuroImage* 24, 862–873.

Delgado, M.R., Frank, R.H., and Phelps, E.A. (2005b). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8, 1611–1618.

Di Chiara, G., Loddo, P., and Tanda, G. (1999). Reciprocal changes in prefrontal and limbic dopamine responsiveness to aversive and rewarding stimuli after chronic mild stress: implications for the psychobiology of depression. *Biol. Psych.* 46, 1624–1633.

Dreher, J.C., Kohn, P., and Berman, K.F. (2006). Neural coding of distinct statistical properties of reward information. *Cerebral Cortex* 16, 561–573.

Drevets, W.C., Gautier, C., Price, J.C. *et al.* (2001). Amphetamine-induced dopamine release in human ventral striatum correlates with euphoria. *Biol. Psych.* 49, 81–96.

Elliott, R., Baker, S.C., Rogers, R.D. *et al.* (1997). Prefrontal dysfunction in depressed patients performing a complex planning task: a study using positron emission tomography. *Psychol. Med.* 27, 931–942.

Elliott, R., Friston, K.J., and Dolan, R.J. (2000). Dissociable neural responses in human reward systems. *J. Neurosci.* 20, 6159–6165.

Finger, S. (1994). *Origins of Neuroscience: A History of Explorations into Brain Function*. Oxford: Oxford University Press.

Friedman, D.P., Aggleton, J.P., and Saunders, R.C. (2002). Comparison of hippocampal, amygdala, and perirhinal projections to the nucleus accumbens: combined anterograde and retrograde tracing study in the Macaque brain. *J. Comp. Neurol.* 450, 345–365.

Gerfen, C.R. and Wilson, C.J. (1996). The basal ganglia. In: A. Bjorklund and T. Hokfelt (eds), *Handbook of Chemical Neuroanatomy*. London: Elsevier Science, pp. 371–468.

Glimcher, P. (2003). *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics*. Cambridge, MA: MIT Press.

Groenewegen, H.J., Wright, C.I., Beijer, A.V.J., and Voorn, P. (1999). Convergence and segregation of ventral striatal inputs and outputs. *Ann. NY Acad. Sci.* 877, 49–63.

Haber, S.N. (2003). The primate basal ganglia: parallel and integrative networks. *J. Chem. Neuroanat.* 26, 317–330.

Haber, S.N. and McFarland, N.R. (1999). The concept of the ventral striatum in nonhuman primates. *Ann. NY Acad. Sci.* 377, 33–48.

Haber, S.N., Fudge, J.L., and McFarland, N.R. (2000). Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J. Neurosci.* 20, 2369–2382.

Haruno, M., Kuroda, T., Doya, K. *et al.* (2004). A neural correlate of reward–based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. *J. Neurosci.* 24, 1660–1665.

Hernandez, L. and Hoebel, B.G. (1988). Food reward and cocaine increase extracellular dopamine in the nucleus accumbens as measured by microdialysis. *Life Sciences* 42, 1705–1712.

Herrero, M.-T., Barcia, C., and Navarro, J.M. (2002). Functional anatomy of thalamus and basal ganglia. *Child Nerv. Sys.* 18, 386–404.

Hess, W.R. (1959). *The Functional Organization of the Diencephalon*. New York: Grune and Stratton.

Hikosaka, O., Sakamoto, M., and Usui, S. (1989). Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *J. Neurophysiol.* 61, 814–832.

Holt, D.J., Graybiel, A.M., and Saper, C.B. (1997). Neurochemical architecture of the striatum. *J. Comp. Neurol.* 384, 1–25.

Ikemoto, S. and Panksepp, J. (1999). The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Res. Rev.* 31, 6–41.

Imperato, A., Angelucci, L., Casolini, P. *et al*. (1992). Repeated stressful experiences differently affect limbic dopamine release during and following stress. *Brain Res.* 577, 194–199.

Jacobs, B.L. and Azmitia, E.C. (1992). Structure and function of the brain serotonin system. *Physiol. Rev.* 72, 165–229.

Jasper, H. and Penfield, W. (1954). *Epilepsy and the Functional Anatomy of the Human Brain*. New York: Little, Brown and Co.

Jeanblanc, J., Hoeltzel, A., and Louilot, A. (2002). Dissociation in the involvement of dopaminergic neurons innervating the core and shell subregions of the nucleus accumbens in latent inhibition and affective perception. *Neuroscience* 111, 315–323.

Kable, J.W. and Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.

Kennedy, R.T., Watson, C.J., Haskins, W.E. *et al*. (2002). In vivo neurochemical monitoring by microdialysis and capillary separations. *Curr. Opin. Chem. Biol.* 6, 659–665.

King-Casas, B., Tomlin, D., Anen, C. *et al*. (2005). Getting to know you: reptutation and trust in a two-person economic exchange. *Science* 308, 78–83.

Kissinger, P.T., Hart, J.B., and Adams, R.N. (1973). Voltammetry in brain tissue: a new neurophysiological measurement. *Brain Research* 55, 209–231.

Knutson, B. and Cooper, J.C. (2005). Functional magnetic resonance imaging of reward prediction. *Curr. Opin. Neurol* 18, 411–417.

Knutson, B. and Gibbs, S.E.B. (2007). Linking nucleus accumbens dopamine and blood oxygenation. *Psychopharmacology* 191, 813–822.

Knutson, B., Westdorp, A., Kaiser, E., and Hommer, D. (2000). fMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage* 12, 20–27.

Knutson, B., Adams, C.M., Fong, G.W., and Hommer, D. (2001a). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* 21, RC159.

Knutson, B., Fong, G.W., Adams, C.M. *et al*. (2001b). Dissociation of reward anticipation and outcome with event-related fMRI. *NeuroReport* 12, 3683–3687.

Knutson, B., Fong, G.W., Bennett, S.M. *et al*. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *NeuroImage* 18, 263–272.

Knutson, B., Taylor, J., Kaufman, M. *et al*. (2005). Distributed neural representation of expected value. *J. Neurosci.* 25, 4806–4812.

Knutson, B., Rick, S., Wimmer, G.E. *et al*. (2007). Neural predictors of purchases. *Neuron* 53, 147–156.

Knutson, B., Wimmer, G.E., Kuhnen, C.M., and Winkielman, P. (2008). Nucleus accumbens activation mediates the influence of reward cues on financial risk taking. *NeuroReport* 19, 509–513.

Koepp, M.J., Gunn, R.N., Lawrence, A.D. *et al*. (1998). Evidence for striatal dopamine release during a video game. *Nature* 393, 266–268.

Kuhnen, C.M. and Knutson, B. (2005). The neural basis of financial risk–taking. *Neuron* 47, 763–770.

Kunig, G., Leenders, K.L., Martin-Soelch, C. *et al*. (2000). Reduced reward processing in the brains of Parkinsonian patients. *NeuroReport* 11, 3681–3687.

LeDoux, J.E. (2000). Emotion circuits in the brain. *Annu. Rev. Neurosci.* 23, 155–184.

Lehericy, S., Ducros, M., van de Moortele, P. *et al*. (2004). Diffusion tensor fiber tracking shows distinct corticostriatal circuits in humans. *Ann. Neurol.* 55, 522–529.

Logothetis, N.K., Pauls, J., Augath, M. *et al*. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150–157.

Lohrenz, T., McCabe, K., Camerer, C.F., and Montague, P.R. (2007). Neural signature of fictive learning signals in a sequential investment task. *PNAS* 104, 9493–9498.

Mark, G.P., Blander, D.S., and Hoebel, B.G. (1991). A conditioned stimulus decreases extracellular dopamine in the nucleus accumbens after the development of a learned taste aversion. *Brain Res.* 551, 308–310.

Martel, P. and Fantino, M. (1996). Mesolimbic dopaminergic system activity as a function of food reward: a microdialysis study. *Pharmacol. Biochem. Behav.* 53, 221–226.

Martin, G.E. and Myers, R.D. (1976). Dopamine efflux from the brain stem of the rat during feeding, drinking, and lever-pressing for food. *Pharmacol. Biochem. Behav.* 4, 551–560.

Mas, M., Fumero, B., and Gonzalez-Mora, J.L. (1995). Voltammetric and microdialysis monitoring of brain monoamine neurotransmitter release during sociosexual interactions. *Behav. Brain Res.* 71, 69–79.

Mawlawi, O., Martinez, D., Slifstein, M. *et al*. (2001). Imaging human mesolimbic dopamine transmission with positron emission tomography: I. Accuracy and precision of D2 receptor parametric measurements in ventral striatum. *J. Cereb. Blood Flow Metab.* 21, 1034–1057.

McClure, S.M., Laibson, D.I., Loewenstein, G.F., and Cohen, J.D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507.

McCullough, L.D. and Salamone, J.D. (1992). Involvement of nucleus accumbens dopamine in the motor activity induced by periodic food persentation: a microdialysis and behavioral study. *Brain Res.* 592, 29–36.

McCullough, L.D., Sokolowski, J.D., and Salomone, J.D. (1993). A neurochemical and behavioral investigation of the involvement of nucleus accumbens dopamine in instrumental avoidance. *Neuroscience* 52, 925–929.

Mitchell, J.B. and Gratton, A. (1991). Opioid modulation and sensitization of dopamine release elicited by sexually relevant stimuli: a high speed chronoamperometric study in freely behaving rats. *Brain Res.* 551, 20–27.

Mogenson, G.M., Jones, D.L., and Yim, C.Y. (1980). From motivation to action: Functional interface between the limbic system and the motor system. *Progr. Neurobiol.* 14, 69–97.

Nakamura, K. and Hikosaka, O. (2006). Facilitation of saccadic eye movements by postsaccadic electrical stimulation in the primate caudate. *J. Neurosci.* 26, 12885–12895.

Newell, A. and Simon, H. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice Hall.

Nieuwenhuys, R. (1985). *Chemoarchitecture of the Brain*. New York, NY: Springer Verlag.

Nieuwenhuis, S., Heslenfeld, D.J., Alting von Geusau, N.J. *et al*. (2005). Activity in human reward-sensitive brain areas is strongly context dependent. *NeuroImage* 25, 1302–1309.

O'Doherty, J.P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* 14, 769–776.

O'Doherty, J., Kringelbach, M.L., Rolls, E.T. *et al*. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat. Neurosci.* 4, 95–102.

O'Doherty, J., Dayan, P., Schultz, J. *et al*. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454.

Olds, M.E. and Fobes, J.L. (1981). The central basis of motivation: intracranial self-stimulation studies. *Annu. Rev. Psychol.* 32, 523–574.

Olds, J. and Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *J. Comp. Physiol. Psychol.* 47, 419–427.

Packard, M.G. and Knowlton, B.J. (2002). Learning and memory functions of the basal ganglia. *Annu. Rev. Neurosci.* 25, 563–593.

Pappata, S., Dehaene, S., Poline, J.B. *et al*. (2002). In vivo detection of striatal dopamine release during reward: a PET study with [(11)C] raclopride and a single dynamic scan approach. *NeuroImage* 16, 1015–1027.

Parkinson, J.A., Olmstead, M.C., Burns, L.H. *et al*. (1999). Dissociation in effects of lesions of the nucleus accumbens core and shell on appetitive Pavlovian approach behavior and the potentiation of conditioned reinforcement and locomotor activity by D-amphetamine. *J. Neurosci.* 19, 2401–2411.

Peters, J.L., Miner, L.H., Michael, A.C., and Sesack, S.R. (2004). Ultrastructure at carbon fiber microelectrode implantation sites after acute voltammetric measurements in the striatum of anesthetized rats. *J. Neurosci. Methods* 137, 9–23.

Pfaus, J.G., Damsma, G., Nomikos, G.G. *et al*. (1990). Sexual behavior enhances central dopamine transmission in the male rat. *Brain Res.* 530, 345–348.

Phillips, A.G., Atkinson, L.J., Blackburn, J.R., and Blaha, C.D. (1993). Increased extracellular dopamine in the nucleus accumbens of the rat elicited by a conditional stimulus for food: an electrochemical study. *Can. J. Physiol. Pharmacol.* 71, 387–393.

Phillips, P.E.M. and Wightman, R.M. (2003). Critical guidelines for validation of the selectivity of in-vivo chemical microsensors. *Trends Anal. Chem.* 22, 509–514.

Phillips, P.E.M., Stuber, G.D., Heien, M.L. *et al*. (2003). Subsecond dopamine release promotes cocaine seeking. *Nature* 422, 614–618.

Poldrack, R.A., Prabhakaran, V., Seger, C.A., and Gabrieli, J.D.E. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology* 13, 564–574.

Preuschoff, K., Bossaerts, P., and Quartz, S.R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51, 381–390.

Radakishun, F.S., van Ree, J.M., and Westerink, B.H.C. (1988). Scheduled eating increases dopamine release increases dopamine release in the nucleus accumbens of food-deprived rats as assessed with on-line brain dialysis. *Neurosci. Letts* 85, 351–356.

Rebec, G.V., Grabner, C.P., Johnson, M. *et al*. (1997). Transient increases in catecholaminergic activity in medial prefrontal cortex and nucleus accumbens shell during novelty. *Neuroscience* 76, 707–714.

Richardson, N.R. and Gratton, A. (1998). Changes in medial prefrontal cortical dopamine levels associated with response-contingent food reward: an electrochemical study in rats. *J. Neurosci.* 18, 9130–9138.

Rilling, J., Gutman, D., Zeh, T. *et al*. (2002). A neural basis for social cooperation. *Neuron* 35, 395–405.

Robbins, T.W., Cador, M., Taylor, J.R., and Everitt, B.J. (1989). Limbic–striatal interactions in reward-related processes. *Neurosci. Biobeh. Rev.* 13, 155–162.

Robinson, D.L., Phillips, P.E.M., Budygin, E.A. *et al*. (2001). Subsecond changes in accumbal dopamine during sexual behavior in male rats. *NeuroReport* 12, 2549–2552.

Robinson, D.L., Heien, M.L., and Wightman, R.M. (2002). Frequency of dopamine concentration transients increases in dorsal and ventral striatum of male rats during introduction of conspecifics. *J. Neurosci.* 22, 10477–10486.

Roitman, M.F., Stuber, G.D., Phillips, P.E.M. *et al*. (2004). Dopamine operates as a subsecond modulator of food seeking. *J. Neurosci.* 24, 1265–1271.

Rolls, E.T. (1994). Neurophysiology and cognitive functions of the striatum. *Revue Neurologie* 150, 648–660.

Rumelhart, D.E. and McClelland, J.L. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press.

Sabol, K.E., Richards, J.B., and Freed, C.R. (1990). In vivo dialysis measurements of dopamine and DOPAC in rats trained to turn on a circular treadmill. *Pharmacol. Biochem. Behav.* 36, 21–28.

Salamone, J.D. (1996). The behavioral neurochemistry of motivation: methodological and conceptual issues in studies of the dynamic activity of nucleus accumbens dopamine. *J. Neurosci. Methods* 64, 137–149.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Sorg, B.A. and Kalivas, P.W. (1991). Effects of cocaine and footshock stress on extracellular dopamine levels in the ventral striatum. *Brain Res.* 559, 29–36.

Talairach, J. and Tournoux, P. (1988). *Co-planar Stereotactic Atlas of the Human Brain*. New York, NY: Thieme.

Thut, G., Schultz, W., Roelcke, U. *et al*. (1997). Activation of the human brain by monetary reward. *NeuroReport* 8, 1225–1228.

Tricomi, E.M., Delgado, M.R., and Fiez, J.A. (2004). Modulation of caudate activity by action contingency. *Neuron* 41, 281–292.

Tricomi, E.M., Delgado, M.R., McCandliss, B.D. *et al*. (2006). Performance feedback drives caudate activation in a phonological learning task. *J. Cogn. Neurosci.* 18, 1029–1043.

Voorn, P., Vanderschuren, L.J., Groenewegen, H.J. *et al*. (2004). Putting a spin on the dorsal–ventral divide of the striatum. *Trends Neurosci.* 27, 468–474.

Westerink, B.H.C. (1995). Brain microdialysis and its application for the study of animal behaviour. *Behav. Brain Res.* 70, 103–124.

Wightman, R.M. and Robinson, D.L. (2002). Transient changes in mesolimbic dopamine and their association with "reward". *J. Neurochem.* 82, 721–735.

Wilkinson, L.S., Humby, T., Killcross, A.S. *et al*. (1998). Dissociations in dopamine release in medial prefrontal cortex and ventral striatum during the acquisition and extinction of classical aversive conditioning in the rat. *Eur. J. Neurosci.* 10, 1019–1026.

Wittmann, B.C., Schott, B.H., Guderian, S. *et al*. (2005). Reward-related FMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron* 45, 459–467.

Yacubian, J., Glascher, J., Schroeder, K. *et al*. (2006). Dissociable systems for gain-and loss-related value predictions and errors of prediction in the human brain. *J. Neurosci.* 26, 9530–9537.

Yin, H.H., Ostlund, S.B., Knowlton, B.J., and Balleine, B.W. (2005). The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* 22, 513–523.

Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2006). Inactivation of the dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behav. Brain Res.* 166, 189–196.

Yoshida, M., Yokoo, H., Mizoguchi, K. *et al*. (1992). Eating and drinking cause increased dopamine release in the nucleus accumbens and ventral tegmental area in the rat: measurement by in vivo microdialysis. *Neurosci. Letts* 139, 73–76.

Young, A.M., Joseph, M.H., and Gray, J.A. (1992). Increased dopamine release in vivo in nucleus accumbens and caudate nucleus of the rat during drinking: a microdialysis study. *Neuroscience* 48, 871–876.

Young, A.M.J., Joseph, M.H., and Gray, J.A. (1993). Latent inhibition of conditioned dopamine release in rat nucleus accumbens. *Neuroscience* 54, 5–9.

Zahm, D.S. and Brog, J.S. (1992). On the significance of subterritories in the "accumbens" part of the rat ventral striatum. *Neuroscience* 50, 751–767.

Zald, D.H., Boileau, I., El-Dearedy, W. *et al*. (2004). Dopamine transmission in the human striatum during monetary reward tasks. *J. Neurosci.* 24, 4105–4112.

Zetterstrom, T., Herrera-Marschitz, M., and Ungerstedt, U. (1981). Simultaneous estimation of dopamine release and rotational beahvior induced by D-amphetamine in 6-OHDA denervated rats. *Neurosci. Letts* 7(Suppl), S27.

CHAPTER

26

The Basal Ganglia and the Encoding of Value

Kenji Doya and Minoru Kimura

OUTLINE

INTRODUCTION

What mechanism of the brain underlies our flexible learning of choice behaviors? According to the theory of reinforcement learning (Sutton and Barto, 1998), an adaptive agent learns behaviors by repeating the following three steps:

1. Predicting the *value* of each action candidate, or option[1]
2. Selecting an action with the highest predicted value
3. Updating the value of the action by the difference between the prediction and the actual outcome.

How can these steps of valuation, action selection, and prediction-error based learning be realized in the brain? Reward-predictive neuron firing has been reported from variety of cortical and subcortical areas, such as the orbitofrontal cortex (Schultz and Dickinson, 2000), the prefrontal cortex (Watanabe, 1996), the parietal cortex (Dorris and Glimcher, 2004; Sugrue *et al*., 2004), and the striatum (Kawagoe *et al*., 1998). Neural firing proportional to reward-prediction error has been reported for midbrain dopamine neurons (Schultz *et al*., 1997; Satoh *et al*., 2003; Bayer and Glimcher, 2005). Functional brain-imaging experiments also report reward-predictive and

---

[1] While the term *option* is commonly used for choice candidates in economics, the term *action* is generally used in reinforcement learning literature, while option often means a higher-level choice of a series of actions.

prediction-error related activities in these areas (O'Doherty *et al.*, 2004; Haruno and Kawato, 2006). The striatum receives a rich dopaminergic input and the cortico-striatal synapses show dopamine-dependent plasticity (Wickens *et al.*, 1996; Reynolds *et al.*, 2001). Based on these observations, we have proposed a schematic framework of how the above three steps of computation can be realized in the cortico-basal ganglia circuit, as depicted in Figure 26.1 (Doya, 2000, 2002, 2007; Daw and Doya, 2006).

Here we further address the following questions:

1. How are values for different actions evaluated in different timescales represented in the striatum?
2. What is the role of the thalamic neurons in action selection?
3. How do the dopamine neurons guide learning of reward prediction?

Based on our own findings and those of others, we propose a more elaborate model of value estimation, action selection and learning in the cortico-basal ganglia circuit.

ACTION-VALUE CODING IN STRIATAL NEURONS

The most popular method in reinforcement learning is to learn the *action value*

$$Q(a) = E[r|a] \tag{26.1}$$

which represents the reward r (amount × probability) expected when taking an action a. If the action values are learned for all possible actions, the obvious choice is to take the action a that gives the largest action value $Q(a)$. Does the brain use such a method and, if so, where in the brain are such action values represented?

In saccadic eye-movement experiments with monkeys, Hikosaka and colleagues showed that firing of striatal projection neurons before saccades was modulated by the amount of reward the monkey had learned it would receive following a successful saccade to a given visual target (Kawagoe *et al.*, 1998). Such action-specific, reward-predictive activities are reminiscent of the action value above. However, their saccade task did not involve a choice between multiple action candidates; only one target was presented in each trial. In order to test how striatal neurons are involved in a reward-based free-choice situation, we performed recording experiments in which a monkey chose one of two possible actions based on the varied probability of liquid rewards (Samejima *et al.*, 2005).

Two macaque monkeys performed a reward-based, free-choice task of turning a handle to the left or the right. The monkeys held the handle in the center position, using their left hand, for a delay period of 1s, and then turned the handle to either the left (a = L) or the right (a = R). An LED on the selected side was illuminated in either green, notifying a large reward (0.2 ml water), or red, notifying a small reward (0.07 ml water). The probabilities of receiving a large reward

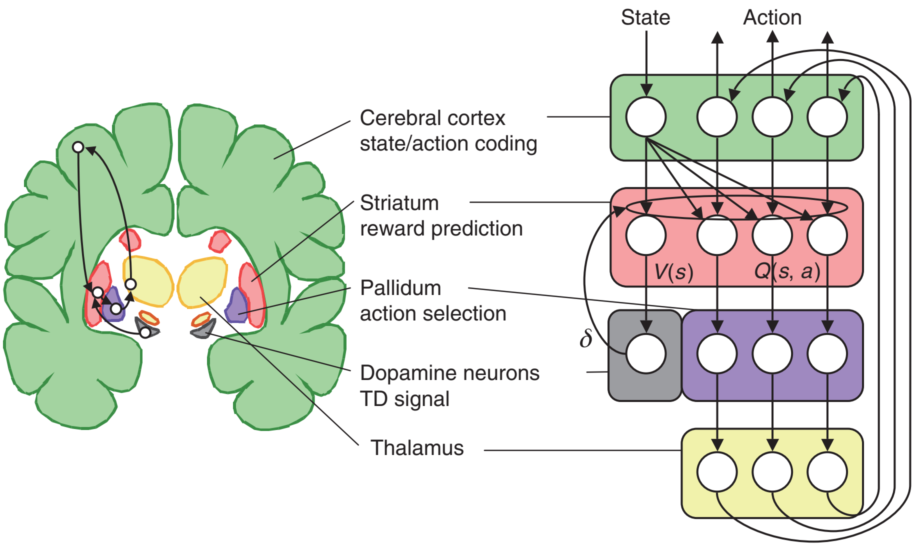


FIGURE 26.1 A schematic model of implementation of reinforcement learning in the cortico-basal ganglia circuit (Doya, 1999, 2000, 2007). Based on the state representation in the cortex, the striatum learns the state value and action values. The state-value coding striatal neurons project to dopamine neurons, which send the TD signal back to the striatum. The outputs of action-value coding striatal neurons channel through the globus gallidus and the thalamus, where stochastic action selection may be realized (Reproduced from Doya, 2007).

after turns to the left and to the right were fixed during a block of 30 to 150 trials, and varied between five different trial blocks. In the 90–50 block, for example, the probability of a large reward for a left turn was 90%, and for a right turn was 50%. In this case, by taking the small reward as the baseline ($r = 0$) and the large reward as unity ($r = 1$), the left action-value Q_L was 0.9 and the right action-value Q_R was 0.5. Four asymmetrically rewarded blocks, 90–50, 50–90, 50–10, and 10–50, and one symmetrically rewarded block, 50–50, were used. An important feature of this block design is that the neuronal activity related to action value can be dissociated from that related to action choice. Suppose a monkey chooses the action with the higher action-value after sufficient learning in a given trial block. While the monkey would be expected to prefer a left turn in both the 90–50 and 50–10 blocks, the action value Q_L for the left turn differs, at 0.9 and 0.5 respectively. Conversely, in the 90–50 and 10–50 blocks, while the monkey's choice behavior should be the opposite (i.e. a left turn in the former and a right turn in the latter), the action value Q_R remains the same at 0.5. Through analysis of choice and reward sequences of two monkeys, we verified that the action-value based model could predict their choice behavior very well.

We recorded 504 striatal projection neurons in the right putamen and caudate nucleus of the two monkeys. Here, we focus on the 142 neurons that displayed increased discharges during at least one task event, and had discharge rates higher than 1 spike/s during the delay pre-movement period. We compared the average discharge rates during the pre-movement delay period from two asymmetrically rewarded blocks. The comparison was based on the trials after the monkeys' choices had reached a "stationary phase" during each block, when the choice probability was biased toward the action with a higher reward probability in more than 70% of trials. In one-half of the neurons (72/142 in two monkeys), activity was modulated by either Q_L or Q_R. For example, the delay period discharge rate of some neurons was significantly higher in the 90–50 block than in the 10–50 block, but was not significantly different between the 50–10 and 50–90 blocks, for which the preferred actions differed. These neurons thus appear to encode the left action-value, Q_L, but not the action or choice itself. Other neurons showed significantly different discharge rates between the 50–10 block and the 50–90 block, but there was no significant difference between the 10–50 and 90–50 blocks. The firing rates of these neurons appear to code the right action-value, Q_R.

Through a multiple regression analysis of neuronal discharge rates with Q_L and Q_R as regressors, we found 24 (17%) "Q_L-type" neurons that had a significant regression coefficient to Q_L (t-test, $P < 0.05$) but not to Q_R, 31 (22%) "Q_R-type" neurons that correlated to Q_R but not to Q_L, and 16 (11%) differential action-value ("Δ-Q-type") neurons that correlated with the difference between Q_L and Q_R. One neuron (<1%) had significant coefficients that correlated to both Q_L and Q_R with the same sign. There were 18 motor-related ("m-type") neurons that had significant t-values only for the action being chosen. However, the discharge rates of most action-value neurons (19/24 in the Q_L-type, 24/31 in the Q_R-type) were not correlated significantly with the action being chosen. We concluded that, during a delay period before action execution, more than one-third of striate projection neurons examined (43/142) encoded action values, and that 60% (43/72) of all the reward value-sensitive neurons were action-value neurons.

Action-value coding in the striatum may be a core feature of information-processing in the basal ganglia. The striatum is the primary target of dopaminergic signals which regulate the plasticity (change in the strength) of cortico-striatal synaptic transmission (Calabresi *et al.*, 1996; Reynolds *et al.*, 2001), conveying signals of actions and cognition. Thus, the striatum may be the locus where reward value is first encoded in the brain.

A recent experiment compared action-value and chosen-value representations in the dorsal striatum and the internal globus pallidus (Pasquereau *et al.*, 2007). These authors found action-value coding in both the striatum and the globus pallidus, but the number of neurons encoding the chosen action increased in the globus pallidus as each trial progressed toward the time of action initiation. This finding, along with our finding of relatively few action-coding neurons in the striatum, favors the view that action selection is realized downstream of the basal ganglia (Doya, 2000; Watanabe *et al.*, 2003) rather than in the striatum itself (Houk *et al.*, 1995; O'Doherty *et al.*, 2004). Further studies on the neuronal activity before and after action selection from different stages of the cortico-basal ganglia loop are necessary to clarify where and how action selection is realized.

SHORT- AND LONG-TERM REWARD PREDICTION IN THE STRIATUM

In the previous section, we considered the case where a reward is given immediately after each action choice. In a more general scenario, an action can result in a reward after various delays, and thus it is not necessarily obvious which of the previous actions is responsible for a given reward. A common way of resolving such a *temporal credit assignment problem* is to learn to predict the cumulative future rewards in the form of the *state value*

$$V(s) = E[r(t) + \gamma r(t+1) + \gamma^2 r(t+2) + \ldots | s(t) = s] \tag{26.2}$$

where state s is defined by a sensory cue (or any other information that is useful for predicting the future outcome), E[] indicates the mean expected value under the current policy (state-to-action mapping), and γ is a parameter called the *temporal discount factor*. The state value is a measure of the long-term goodness of the given state s under the current policy; thus the increase or decrease in $V(s)$ can be regarded as a virtual reward signal. More precisely, the inconsistency in the prediction

$$\delta(t) = r(t) + \gamma V(s(t+1)) - V(s(t)) \tag{26.3}$$

is termed the *temporal-difference (TD) error*, and this can be utilized as the effective reward signal that takes into account the delayed rewards (Sutton and Barto, 1998).

The discount factor γ defines the temporal focus of this cumulative reward prediction: if $\gamma = 0$, only the immediate reward $r(t)$ is considered; if γ is set close to 1, long-delayed rewards are also taken into account. In essence, γ controls the temporal horizon of future reward estimation. The temporal discount factor γ is a critical parameter that determines the character of learned behaviors (Doya, 2002) – for example, a low setting of γ can lead to short-sighted, impulsive behaviors. In order to understand the brain's mechanism for reward prediction at different timescales, and its potential mechanism of regulation, we performed an fMRI experiment in which subjects learned to take small losses in order to acquire subsequent large rewards (Tanaka *et al*., 2004).

In the *Markov decision task* (Figure 26.2a), one of three states is presented to the subject visually, using three different figures, and the subject selects one of two actions by pressing one of two buttons. For each state, the subject's action affects not only the reward given immediately but also the state subsequently presented. In the SHORT condition, action a_1 gives a small positive reward $+r_1$ (20 yen average) and action a_2 gives a small negative reward $-r_1$ (-20 yen average) at all three states. The optimal behavior for maximizing the total outcomes is to collect small positive rewards by taking action a_1 at each state. In the LONG condition, while action a_2 at state s_3 gives a big bonus $+r_2$ (100 yen average), action a_1 at state s_1 results in a big loss $-r_2$ (-100 yen average). The optimal behavior is to receive small negative rewards to obtain a large positive reward by taking action a_2 at each state – the opposite to the optimal behavior in the SHORT condition. Thus, in the LONG condition, the subject has to select an action by taking into account both the immediate reward and the future reward expected from the subsequent state, while in the SHORT condition the subject needs to consider only the immediate outcome. In order to remove those vision and motor-related brain activities independent of reward processing, we introduced the NO reward condition, in which the reward was always zero and the subject was free to choose any button.

We first performed block-design analyses to assess the brain areas specifically involved in short- and long-term reward prediction. In the statistical comparison of brain activities during the SHORT vs NO conditions, a significant increase in activity was observed in the lateral orbitofrontal cortex (OFC), the insula, and the occipitotemporal area (OTA), as well as in the striatum, globus pallidus (GP), and medial cerebellum. These areas may be involved in reward prediction that only takes into account immediate outcome. In the LONG vs SHORT contrast, a robust increase in activity was observed in the ventrolateral PFC (VLPFC), insula, dorsolateral prefrontal cortex (DLPFC), dorsal premotor cortex (PMd), inferior parietal cortex (IPC), striatum, globus pallidus, dorsal raphe nucleus, lateral cerebellum, posterior cingulate cortex, and subthalamic nucleus. These areas are specifically involved in decision making based on prediction of reward in multiple steps in the future, which was specifically required in the LONG condition but not in the SHORT condition. The results of these block-design analyses suggest differential involvement of brain areas in predicting immediate and future rewards. These results are also consistent with a more recent fMRI study using an inter-temporal choice task, which found activities in the lateral prefrontal and parietal cortex for delayed reward choice (McClure *et al*., 2004).

In order to further clarify the brain structures specific to reward prediction at different timescales, we estimated how much reward the subjects should have predicted on the basis of their behavioral data. We then used these trial-by-trial predictions to construct time-courses as the explanatory variables for a regression analysis. Specifically, we estimated the time-courses of reward prediction $V(s(t))$ and prediction error $\delta(t)$, as defined in equations (26.2) and (26.3), from each subject's performance data. In our Markov decision task, the minimum value of γ needed to find the optimal action in the LONG condition was 0.36, while any small value of γ was sufficient in the

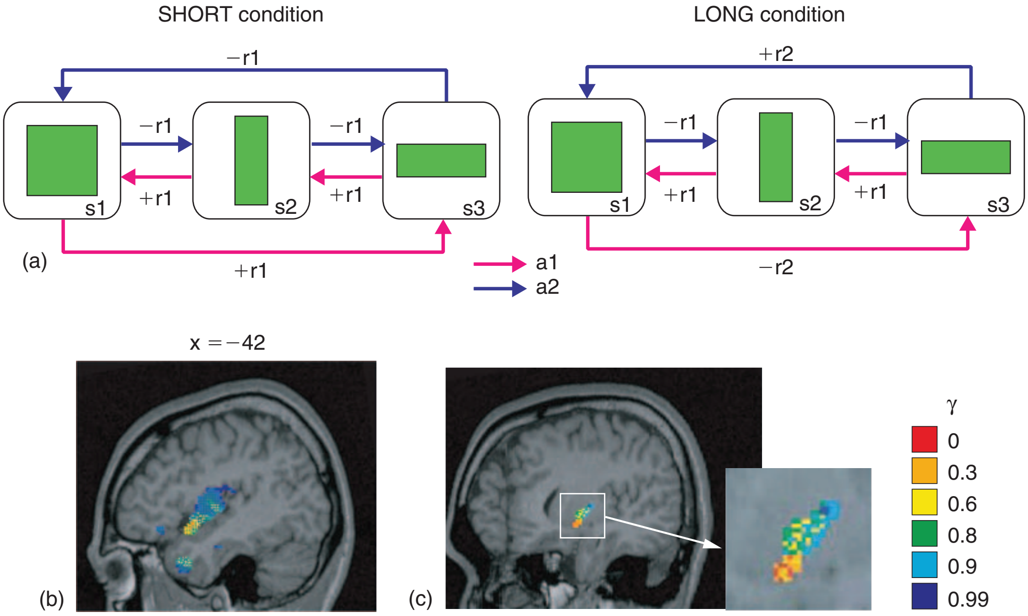


FIGURE 26.2 Experimental design of the Markov decision task. (a) At the beginning of each trial block, the condition is informed by displaying its character (e.g. "SHORT condition"). A fixation point is presented on the screen, and after 2 seconds one of three figures (square, vertical rectangle, or horizontal rectangle) is presented. As the fixation point vanishes after 1 s, the subject presses either the right or left button within 1 s. After a short delay (1 s), a reward for the current action is presented by a number, and the past cumulative reward is shown by a bar graph. A single trial takes 6 seconds. The rules of the reward and state transition for actions a_1 (red arrow) and a_2 (blue arrow) are shown in the SHORT and LONG conditions. The small reward r_1 is 10, 20, or 30 yen, with equal probability, and the large reward r_2 is 90, 100, or 110 yen. The rule of state transition is the same for all conditions; $s_3 \rightarrow s_2 \rightarrow s_1 \rightarrow s_3 \ldots$ for action a_1, and $s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_1 \ldots$ for action a_2. Although the optimal behaviors are opposing (action a_1 in the SHORT condition and action a_2 in the LONG condition), the expected cumulative reward during one cycle of the optimal behavior is 60 yen in both the SHORT ($+20 \times 3$) and the LONG ($-20 - 20 + 100$) conditions. (b) Voxels with a significant correlation with reward prediction $V(t)$ in the insular cortex (height threshold of $P < 0.001$, uncorrected; extent threshold 4 voxels). (c) Voxels with a significant correlation with reward-prediction error $\delta(t)$ in the striatum. Different colors are used for different settings of the discount factor: $\gamma = 0$, red; $\gamma = 0.3$, orange; $\gamma = 0.6$, yellow; $\gamma = 0.8$, green; $\gamma = 0.9$, cyan; $\gamma = 0.99$, blue). Note the ventromedial to dorsolateral gradient with the increase in γ in both the insula and the striatum. (Reproduced from Tanaka *et al*., 2004).

SHORT condition. We hypothesized that different parts of the brain are specialized for reward prediction at different timescales. Accordingly, we estimated $V(s(t))$ and $\delta(t)$ using six different levels of the discount factor ($\gamma = 0$, 0.3, 0.6, 0.8, 0.9, and 0.99), and searched for voxels that had significantly correlated time-courses to those explanatory variables.

We observed a significant correlation with reward prediction $V(s(t))$ in the medial prefrontal cortex (mPFC) and the bilateral insula (Figure 26.2b), the left hippocampus, and the left temporal pole (the foremost part of the temporal lobe). The activities of the medial prefrontal cortex (mPFC), temporal pole, and hippocampus correlated with reward prediction $V(s(t))$ with a longer timescale ($\gamma = 0.6$). Furthermore, as shown in Figure 26.2b, using graded colors for different discount factors γ (red for $\gamma = 0$, blue for $\gamma = 0.99$), we found a graded map of activities for reward prediction at different timescales in the insula. While the activity in the ventromedial part correlated with reward prediction at a shorter timescale, the activity of the dorsolateral part correlated with reward prediction at a longer timescale. We also found significant correlation with reward-prediction error $\delta(t)$ with a wide range of timescales in the striatum (Figure 26.2c). Again, a graded map was produced, which had a short timescale in the ventromedial part and a long timescale in the dorsolateral part.

The results of the block-design and performance-based regressor analyses suggest differential involvement of brain areas in action learning by prediction of rewards at different timescales. In the insula and the anterior striatum, activities were found both in block-design and performance-based regression analyses. The vertical anatomical shifts in the activated locus in the SHORT vs NO and LONG vs SHORT contrasts in each area are consistent with the ventro-dorsal maps of the discount factor γ found in the performance-based regression analysis. Correlation of the striatal activity with reward-prediction error $\delta(t)$ could be due to dopamine-dependent plasticity of cortico-striatal synapses (Reynolds and Wickens, 2002).

In summary, compared with the control task in which subjects simply learned to acquire immediate positive rewards, we found enhanced activity in the prefrontal, premotor, and parietal cortices, as well as in the dorsal striatum, lateral cerebellum, and the midbrain including the dorsal raphe nucleus. By reinforcement-learning model-based analysis using multiple temporal discounting parameters, we found a ventral-to-dorsal map of short-to-long timescales of reward prediction in the striatum and the insular cortex.

CENTROMEDIAN THALAMIC NEURONS

As we mentioned earlier in the chapter, the site and the mechanism of value-based action selection are still to be investigated. Thus we have examined the roles of the thalamo-striate projection in reward value-based action selection in the basal ganglia. The centromedian/parafascicular (CM/PF) complex of the thalamus (Steriade *et al.*, 1997) has received little attention in the studies of action and cognition; however, its outputs direct mostly to the putamen and caudate nucleus as well as to the medial frontal cortex, and it receives topographically organized inputs from the output stations of the basal ganglia as well as from the reticular formation, superior colliculus, and pedunculo-pontine tegmental nucleus (Groenewegen and Berendse, 1994; Steriade *et al.*, 1997; Matsumoto *et al.*, 2001; Takada *et al.*, 2001; Smith *et al.*, 2004). There are two representative types of neurons in the CM/PF complex: one exhibits an increase in discharges after visual, auditory, and somatosensory stimuli at very short latency (SLF), while the other exhibits facilitatory responses at long latency (>200 ms, LLF) (Minamimoto and Kimura, 2002; Minamimoto *et al.*, 2005). SLF neurons are mostly located in the PF, while LLF neurons are mostly located in the CM (Matsumoto *et al.*, 2001).

We recorded from LLF neurons in the CM in an asymmetrically rewarded GO/NO-GO task in which two kinds of visual stimuli were presented, one for the GO response and the other for the NO-GO response (Minamimoto *et al.*, 2005). Performance of the requested action, whether GO or NO-GO, was rewarded by a large amount of water, while performance of the other action was rewarded with a small amount of water. The action–outcome association was then altered in the next block. Monkeys performed the large-reward GO trials with shorter reaction times than they did the small-reward trials. This indicated that the monkeys assigned higher values to the large-reward actions and prepared for them (biased toward the actions). In addition, the rate of error trials, such as too long reaction times or initiation of incorrect actions, was higher in the small-reward trials.

Remarkably, the majority of the LLF neurons exhibited burst discharges selectively after the visual cue indicated that the small-reward option would be required, while they showed very little activity after of the visual cue indicated that the large-reward action could be required. This was true regardless of whether the action required was GO or NO-GO. The magnitude of the activity following small-reward action cues became larger, across trials, when the probability of a large-reward action cue increased. Thus, the critical nature of the activity was its specificity to the small-reward option among available actions when subjects were preparing to choose a large-reward option. These results raised the question of whether artificial activation of CM after a GO request could trigger the complementary process to the GO-action bias. Indeed, we found that electrical stimulation of CM after a large-reward GO request significantly slowed down the behavioral response, just as after a GO small-reward request. These results suggest a specific participation of CM in abolishing bias towards the large-reward option, and in pursuing the small-reward action, which are complementary processes to the response bias.

In this experimental situation, subjects had to complete small-reward trials in order to move on, and to obtain large rewards in later trials. Thus, when a small-reward instruction occurred a motivational conflict arose, because subjects had to perform voluntarily an option they did not want to choose. Response bias is not responsible for this situation, but another mechanism is necessary to cancel any premature motor programming through inhibiting the processes of response bias and reinforcing lower-valued but required action.

A recent study supported the involvement of the striato-pallidal system in this process by demonstrating that neurons encoding the values of chosen actions are more dominant and prevalent in the globus pallidus than in the striatum, where neurons dominantly encoding the expected reward probabilities of specific options of action (action values) are prevalent, irrespective of whether the action is to be selected or not (Pasquereau *et al.*, 2007). A subset of neurons in the medial frontal cortex was selectively activated when saccadic eye movements which had been repeatedly triggered based on some rule to one of two previously rewarded directions were switched to the other direction after noticing reversal of action–outcome associations (Isoda and Hikosaka, 2007). The neuronal activation began and terminated mostly before the movement onset in trials in which the monkey was asked to switch direction. Neurons in the cingulate motor area were activated selectively when subjects made switches from previously rewarded (but currently unrewarded) actions to previously unrewarded, currently rewarded actions. Further, this action switch was impaired by chemical inactivation of the cingulate motor area (Shima and Tanji, 1998). Therefore, the next critical issue for the studies of reward value-based decision making, action selection, and monitoring and evaluation of ongoing or performed actions is to understand the roles played specifically by the striatum, striato-pallidal system, thalamo-striatal system, and medial frontal cortical areas.

MOTIVATION AND OUTCOME CODING IN DOPAMINE NEURONS

Dopamine (DA) neurons in the midbrain have been shown to encode prediction errors of probability and/or of magnitude of reward (Schultz, 1998; Satoh *et al.*, 2003; Morris *et al.*, 2004; Nakahara *et al.*, 2004) or timing of expected reward (Bayer and Glimcher, 2005) through single-neuron recording experiments in monkeys performing behavioral tasks. On the other hand, a substantial body of evidence suggests involvement of DA systems in the processes of motivation (Robbins and Everitt, 1996; Koepp *et al.*, 1998; Salamone and Correa, 2002; Wise, 2002), and in switching attentional and behavioral selections to salient stimuli that underlie associative learning (Spanagel and Weiss, 1999; Redgrave *et al.*, 1999). It has been well documented that DA neurons show phasic activations with a wide variety of salient stimuli, including novel and high-intensity stimuli (Jacobs, 1986; Schultz and Romo, 1987; Ljungberg *et al.*, 1992; Horvitz *et al.*, 1997). Why do DA neurons encode reward expectation errors and motivation, and how are these signals integrated with the processes of decision making and learning? We addressed these issues by examining the activity of DA neurons of monkeys that made a series of behavioral decisions based on trial-specific reward expectations (Satoh *et al.*, 2003).

In a multi-step decision and action selection task, Japanese Monkeys depressed a start button with their hand after the button was illuminated. Three target buttons were then simultaneously switched on, and the monkeys released the start button and depressed one of the illuminated target buttons. If an incorrect button was depressed, a low-frequency acoustic tone occurred, and the monkeys chose one of the remaining two buttons in the next trial. If the correct button was depressed, a high-frequency tone occurred, and a small amount of reward water was delivered. The high- and low-tone sounds served as positive and negative reinforcers, respectively, following the behavioral decisions. Once the monkeys had identified the correct button, the same button was depressed correctly button in all subsequent trials. The monkeys received a reward three times, by selecting the same button during three consecutive trials. The trials in a single block were therefore divided into two time periods: the trial-and-error epoch and the repetition epoch. Five types of trials occurred – trials in which the monkeys chose the correct button at the first, second, or third choice in a single block (N1, N2, and N3, respectively) during the trial-and-error epoch, and at the first and the second trials during the repetition epoch (R1 and R2, respectively). Average probabilities of correct choices in N1, N2, N3, R1, and R2 trials were 20%, 50%, 85%, 93%, and 95%, respectively.

Following a few months of trials, the monkeys were found to be performing N1 trials with the longest reaction times and R2 trials with the shortest reaction times, among the five types of trials, after the start button had been illuminated. This suggested that the monkeys had developed trial type-specific levels of reward expectations. Furthermore, errors of reward prediction could be estimated – 80% in N1, 50% in N2, 15% in N3, 7% in R1 and 5% in R2 trials.

During these trials, DA neurons exhibited a maximal increase of discharge rate above their background level of 4–5 spikes/s after a positive reinforcer during N1 trials with a reward probability of 20%. Responses during N2 and N3 trials gradually decreased, and those during R1 and R2 trials were so small that it was difficult to detect increases of discharge rates above the background level. It was found that the magnitude of DA neuron responses reflected precisely the errors of reward expectation, as shown by an excellent fit of estimated values of reward expectation errors (REEs) to the DA responses. On the other hand, decrease of discharge rate of DA neurons after negative reinforcers exhibited a similar tendency – small responses in N1 trials and large responses in R1 and R2 trials – but the estimation of DA neuron responses by negative REEs was poor.

About a half of DA neurons also exhibited trial type-dependent activation after appearance of the start cue – the lowest in N1 trials and the highest in N2 or N3 trials. Thus, the dopamine neuron responses to the visual cue instructing the start of each trial had a tendency to be small when the reward probability of the trial was low, but large when the reward probability was high. However, the responses in the trials with the highest reward probability (>97% in R1 and/or

R2 trials) were smaller than those in either N2 or N3 trials. This was not consistent with previous observations of DA neuron responses of monkeys performing classical conditioning (Fiorillo *et al.*, 2003) and instrumental conditioning paradigms (Morris *et al.*, 2004, 2006). Probably, this was because subjects in our multi-step decision and action selection task expected rewards not only immediately after current trials but also after multi-step trials; this was in contrast to the behavioral situations employed in previous studies, in which only reward probabilities at current trials were predicted. On the other hand, the start cue responses could reflect the level of motivation, or willingness, to start an individual trial in order to obtain an expected reward after the trial. Reaction times of subjects in performing a required action have been used as a measure of motivational level at individual trials (Shidara *et al.*, 1998; Watanabe *et al.*, 2001; Kobayashi *et al.*, 2002; Takikawa *et al.*, 2002). We recorded the reaction times, as well as DA neuron discharges, of monkeys from the appearance of the start cue to performance of the action. DA neuron responses to the start cue changed considerably depending on the subject's reaction time. This suggested that DA neuron responses to the start cue are modulated by motivation. Importantly, the magnitude of DA neuron responses to the start cue was positively correlated with the magnitude of responses to the high-frequency tone (positive reinforcer) that occurred after a correct choice was made.

What is the functional role of the dual coding of incentive attribution to the start cue and of REEs in reward-based decision making and learning? One possible and fascinating role is modulation of the effectiveness of REEs as a teaching signal by a motivation. For instance, the rate of learning could be faster when animals are highly motivated because of stronger activation of DA neurons (and thus larger amount of DA release) and slower when they are less motivated, even at identical REEs, as a consequence of an action. This suggests a new and richer model for DA neurons as teaching signals in reinforcement learning than is currently proposed. It is also consistent with the theory of classical conditioning, in which the rate of learning is assumed to be influenced by factors such as attention or motivation (Rescorla and Wagner, 1972; Dickinson, 1980; Niv *et al.*, 2007). From a computational point of view, involvement of motivational processes in instrumental conditioning has recently been emphasized, and a new model of reinforcement learning has been put forward in which DA neurons transmit both reward-expectation error and impact of motivation (Dayan and Balleine, 2002).

CONCLUSION

Figure 26.3 is an augmented schematic diagram incorporating our findings that LLF neurons in CM encode and transmit signals of anti-bias on selection mainly to the striatum, and that DA neurons encode and transmit signals of REEs, reward value, and motivation to the striatum, where adaptive action-value coding occurs based on the DA neuron signals. Reward-predictive and motivational coding of the dopamine neurons, as well as the complementary

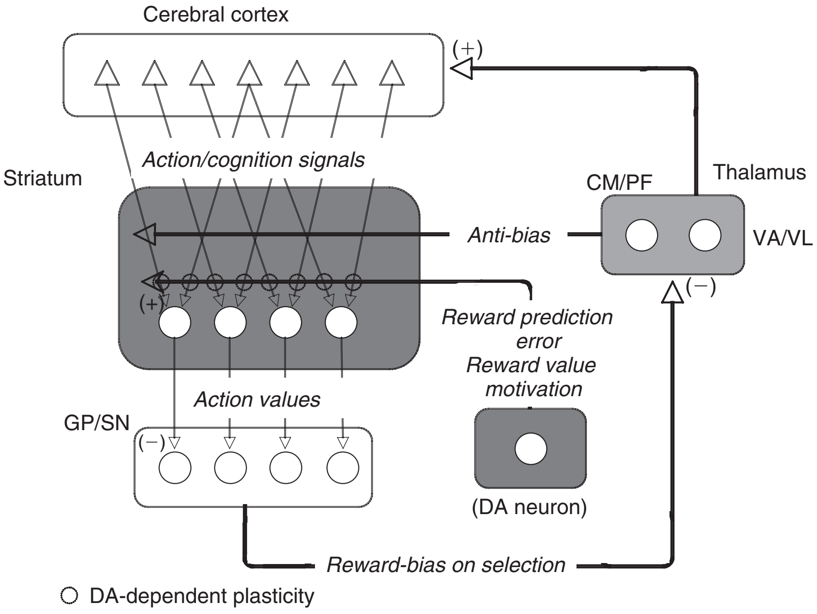


FIGURE 26.3 Schematic diagram showing how the basal ganglia encode reward values of external signals and actions, and how desirable actions are selected. The cortico-basal ganglia loop is composed of the cortico-basal ganglia-thalamo-cortical "external" loop, and the striato-pallido/nigro-thalamo-striate "internal" loop. GP/SN, globus pallidus and substantia nigra; CM/PF, centromedian parafascicular nuclei of intralaminar thalamus.

activities of the CM thalamic neurons, may facilitate the acquisition of action-value coding of the striatal neurons. The mapping of multiple timescales in the ventro-dorsal axis of the cortico-striatal circuit may allow flexible decision making in light of immediate and future costs and benefits.

References

Bayer, H.M. and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141.

Calabresi, P., Pisani, A., Mercuri, N.B., and Bernardi, G. (1996). The corticostriatal projection: from synaptic plasticity to dysfunctions of the basal ganglia. *Trends Neurosci.* 19, 19–24.

Daw, N.D. and Doya, K. (2006). The computational neurobiology of learning and reward. *Curr. Opin. Neurobiol.* 16, 199–204.

Dayan, P. and Balleine, B.W. (2002). Reward, motivation, and reinforcement learning. *Neuron* 36, 285–298.

Dickinson, A. (1980). *Contemporary Animal Learning Theory*. Cambridge: Cambridge University Press.

Dorris, M.C. and Glimcher, P.W. (2004). Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44, 365–378.

Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks* 12, 961–974.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Curr. Opin. Neurobiol.* 10, 732–739.

Doya, K. (2002). Metalearning and neuromodulation. *Neural Networks* 15, 495–506.

Doya, K. (2007). Reinforcement learning: computational theory and biological mechanisms. *HFSP J.* 1, 30–40.

Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902.

Groenewegen, H.J. and Berendse, H.W. (1994). The specificity of the "nonspecific" midline and intralaminar thalamic nuclei. *Trends Neurosci.* 17, 52–57.

Haruno, M. and Kawato, M. (2006). Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops: fMRI examination in stimulus-action-reward association learning. *Neural Networks* 19, 1242–1254.

Horvitz, J.C., Stewart, T., and Jacobs, B.L. (1997). Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. *Brain Res.* 759, 251–258.

Houk, J.C., Adams, J.L., and Barto, A.G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: J.C. Houk, J.L. Davis, and D.G. Beiser (eds), *Models of Information Processing in the Basal Ganglia*. Cambridge, MA: MIT Press, pp. 249–270.

Isoda, M. and Hikosaka, O. (2007). Switching from automatic to controlled action by monkey medial frontal cortex. *Nat. Neurosci.* 10, 240–248.

Jacobs, B.L. (1986). Single unit activity of brain monoamine-containing neurons in freely moving animals. *Ann. NY Acad. Sci.* 473, 70–77.

Kawagoe, R., Takikawa, Y., and Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* 1, 411–416.

Kobayashi, Y., Inoue, Y., Yamamoto, M. *et al*. (2002). Contribution of pedunculopontine tegmental nucleus neurons to performance of visually guided saccade tasks in monkeys. *J. Neurophysiol.* 88, 715–731.

Koepp, M.J., Gunn, R.N., Lawrence, A.D. *et al*. (1998). Evidence for striatal dopamine release during a video game. *Nature* 393, 266–268.

Ljungberg, T., Apicella, P., and Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* 67, 145–163.

Matsumoto, N., Minamimoto, T., Graybiel, A.M., and Kimura, M. (2001). Neurons in the thalamic CM-Pf complex supply striatal neurons with information about behaviorally significant sensory events. *J. Neurophysiol.* 85, 960–976.

McClure, S.M., Laibson, D.I., Loewenstein, G., and Cohen, J.D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507.

Minamimoto, T. and Kimura, M. (2002). Participation of the thalamic CM-Pf complex in attentional orienting. *J. Neurophysiol.* 87, 3090–3101.

Minamimoto, T., Hori, Y., and Kimura, M. (2005). Complementary process to response bias in the centromedian nucleus of the thalamus. *Science* 308, 1798–1801.

Morris, G., Arkadir, D., Nevet, A. *et al*. (2004). Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* 43, 133–143.

Morris, G., Nevet, A., Arkadir, D. *et al*. (2006). Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* 9, 1057–1063.

Nakahara, H., Itoh, H., Kawagoe, R. *et al*. (2004). Dopamine neurons can represent context-dependent prediction error. *Neuron* 41, 269–280.

Niv, Y., Daw, N.D., Joel, D., and Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacol. (Berl)* 191, 507–520.

O'Doherty, J., Dayan, P., Schultz, J. *et al*. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454.

Pasquereau, B., Nadjar, A., Arkadir, D. *et al*. (2007). Shaping of motor responses by incentive values through the basal ganglia. *J. Neurosci.* 27, 1176–1183.

Redgrave, P., Prescott, T.J., and Gurney, K. (1999). Is the short-latency dopamine response too short to signal reward error? *Trends Neurosci.* 22, 146–151.

Rescorla, R.A. and Wagner, A.R. (1972). Current research and theory. In: A.H. Black and W.F. Prokasy (eds), *Classical Conditioning*, Vol. II. New York, NY: Appleton Century Crofts, pp. 64–99.

Reynolds, J.N. and Wickens, J.R. (2002). Dopamine-dependent plasticity of corticostriatal synapses. *Neural Networks* 15, 507–521.

Reynolds, J.N., Hyland, B.I., and Wickens, J.R. (2001). A cellular mechanism of reward-related learning. *Nature* 413, 67–70.

Robbins, T.W. and Everitt, B.J. (1996). Neurobehavioural mechanisms of reward and motivation. *Curr. Opin. Neurobiol.* 6, 228–236.

Salamone, J.D. and Correa, M. (2002). Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behav. Brain Res.* 137, 3–25.

Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340.

Satoh, T., Nakai, S., Sato, T., and Kimura, M. (2003). Correlated coding of motivation and outcome of decision by dopamine neurons. *J. Neurosci.* 23, 9913–9923.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.

Schultz, W. and Dickinson, A. (2000). Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* 23, 473–500.

Schultz, W. and Romo, R. (1987). Responses of nigrostriatal dopamine neurons to high-intensity somatosensory stimulation in the anesthetized monkey. *J. Neurophysiol.* 57, 201–217.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Shidara, M., Aigner, T.G., and Richmond, B.J. (1998). Neuronal signals in the monkey ventral striatum related to progress through a predictable series of trials. *J. Neurosci.* 18, 2613–2625.

Shima, K. and Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. *Science* 282, 1335–1338.

Smith, Y., Raju, D.V., Pare, J.F., and Sidibe, M. (2004). The thalamostriatal system: a highly specific network of the basal ganglia circuitry. *Trends Neurosci.* 27, 520–527.

Spanagel, R. and Weiss, F. (1999). The dopamine hypothesis of reward: past and current status. *Trends Neurosci.* 22, 521–527.

Steriade, M., Jones, E.G., and McCormick, C.D. (1997). *Organisation and Function*. Oxford: Elsevier Science.

Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.

Sutton, R.S. and Barto, A.G. (1998). *Reinforcement Learning*. Cambridge, MA: MIT Press.

Takada, M., Tokuno, H., Hamada, I. *et al*. (2001). Organization of inputs from cingulate motor areas to basal ganglia in macaque monkey. *Eur. J. Neurosci.* 14, 1633–1650.

Takikawa, Y., Kawagoe, R., Itoh, H. *et al*. (2002). Modulation of saccadic eye movements by predicted reward outcome. *Exp. Brain Res.* 142, 284–291.

Tanaka, S.C., Doya, K., Okada, G. *et al*. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat. Neurosci.* 7, 887–893.

Watanabe, K., Lauwereyns, J., and Hikosaka, O. (2003). Neural correlates of rewarded and unrewarded eye movements in the primate caudate nucleus. *J. Neurosci.* 23, 10052–10057.

Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature* 382, 629–632.

Watanabe, M., Cromwell, H.C., Tremblay, L. *et al*. (2001). Behavioral reactions reflecting differential reward expectations in monkeys. *Exp. Brain Res.* 140, 511–518.

Wickens, J.R., Begg, A.J., and Arbuthnott, G.W. (1996). Dopamine reverses the depression of rat corticostriatal synapses which normally follows high-frequency stimulation of cortex *in vitro*. *Neuroscience* 70, 1–5.

Wise, R.A. (2002). Brain reward circuitry: insights from unsensed incentives. *Neuron* 36, 229–240.

PART V

THE NEURAL MECHANISMS FOR CHOICE

CHAPTER

27

The Neural Mechanisms that Underlie Decision Making

Charles R. Gallistel

OUTLINE

INTRODUCTION

The first word that springs to mind in introducing contemporary work on the neurobiology of decision making is *ignoramus*: we do not know the neural basis of economically consequential decision making. However, *ignoramus* does not imply *ignorabimus* – that we *will* not know. Not knowing *now* is not a reason to despair of knowing, but rather the source of our aspiration to know. The chapters in this section are an introduction to contemporary efforts to turn that aspiration into knowledge.

The task is daunting, because it involves the establishment of empirically well-secured hypotheses linking phenomena and entities at different levels of analysis. Three levels of analysis come into play in neuroeconomics: the economic, the psychological, and the neurobiological. The objects of analysis and the terms in which the analyses are conducted at one level have no obvious referents or definitions at the other levels. The economist works at the market level, trying to understand how the decisions made by individuals in a market economy determine quantities such as aggregate supply, aggregate demand, and price. These are definable only by reference to a market economy. They are not the variables that psychologists work with in trying to understand the behavior of human and non-human animals, because they do not operate *within* individual subjects. Neither are they attributes of the behavior of an individual subject. On the other hand, economists do posit variables (for example, subjective utilities) and processes (for example, decision processes) that they suppose do operate within the brains of individual subjects to determine the decisions that collectively create demand, help to determine supply, and establish price. However, these psychological variables and processes have, in their turn, no obvious referents at the neurobiological level. Indeed, in some neurobiological circles there is doubt whether symbolic variables such as subjective values have a role to play in understanding how the brain works (Edelman, 1989).

An illustration of the difficulty of establishing, to general satisfaction, hypotheses that link concepts defined at one level of analysis to concepts defined at a different level is provided by the lengthy history of attempts to link the concept of a nerve impulse to the concept of an action potential. The concept of a nerve impulse is defined by reference to behavioral observations. It has been invoked, at least since the time of Descartes, to explain the fact that pinching the toe gives rise to contractions in muscles far removed from the toe. The concept of an action potential is defined by reference to electrophysiological observations on nerves and muscles, which were first made by du Bois-Reymond in 1848 (see du Bois-Reymond, 1852). He believed that the brief alterations in the electrical potential between the insides of (a bundle of) axons and their outsides that accompanied the behavioral action of a nerve was the physical realization of the nerve impulse. The physical nature of the impulse had heretofore been an object of much speculation, with the vitalists (then still an influential school of thought) doubting that it had a physical realization. Du Bois-Reymond was, of course, right; no one doubts this linkage hypothesis now. However, in 1936, 40 years after Du Bois-Reymond died, the Nobelist A.V. Hill opened a review of experimental work on excitation and accommodation in nerves with the assertion that the physical identity of the nerve impulse remained an unsolved problem (Hill, 1936).

I stress the difficulty of firmly establishing these linkage hypotheses in order to dash one hope that economists might entertain about the fruits of neuroeconomic research: measures of neural signaling activity are unlikely to place hard-to-measure economic decision variables like subjective values on a more secure foundation in our lifetimes. In order for that to happen, we would need to establish a secure hypothesis linking subjective values to measurable neurobiological quantities. We would have to believe, for example, with a high degree of confidence, that subjective values are physically realized by the firing rates of identifiable neurons and that therefore, as Glimcher boldly hypothesizes, "they take as their natural units action potentials per second" (see Chapter 32, page 509)

CAN/SHOULD NEUROBIOLOGICAL DATA CONSTRAIN ECONOMIC THEORIZING?

Closely related to my skepticism regarding the prospects for replacing messy behavioral measures of critical subjective variables with precise neurobiological measures are my doubts about the ability of neurobiological data to constrain economic theorizing. The assumption that neurobiological data can and should constrain both psychological and economic theorizing is widespread and seldom questioned. It is, for example, a foundation for neural net theorizing of all kinds. In Chapter 28, Rangel argues that "A central goal of neuroeconomics is to construct theories of value-based decision making that are neurobiologically grounded," and that, "As in the case of perceptual decision making, a key test of the model is the presence of neurobiological signals that resemble the computations that it makes." And Glimcher argues in Chapter 32 that "The explicit ties to neurobiological data will reveal that only a tiny space in the vast landscape of economic theory can be viewed as compatible with the human neuroarchitecture."

In order for neurobiological data to constrain psychological and economic theorizing, one must be persuaded beyond reasonable doubt that the neurobiological variables that were measured or the neural structures that were described were actually correctly linked to a psychological variable, such as subjective value, in a valid theory of the psychological causation of the behavior. If the neurons whose firing rates are measured are not the neurons whose firing rates physically realize the subjective values in question, then the measures made on them are irrelevant to the psychological or economic theories in which those variables play a role. More radically, if, at the point in the brain where decisions are made, subjective values are not represented by firing rates at all but rather by some other measurable neurobiological variable – for example, the concentrations of intracellular signaling molecules – then any measure of firing rate or of quantities derivative of it, such as BOLD signals, are irrelevant to our understanding of the decision-making process. Moreover, if economically consequential decisions are not made in the cortex but elsewhere – for example, in the hypothalamus – then the architecture of some portion of the cortex is irrelevant. The architecture of the hypothalamus is radically different from the architecture of the cortex, and neuroscientists are more than a little unclear about the functional relevance of the architecture of either one. There is little doubt that their architecture is an important part of their functional capability – why should the brain be different from the liver, or the eye? – but just what their capabilities are and just how the architecture supports them are far from decided questions. In many cases, there are not even any clear hypotheses.

In short, the relevance of neurobiological data to psychological and economic theorizing depends on the strength of the hypotheses that assert that specified neural variables are the physical realization of specified psychological and economic variables. These hypotheses are among the most difficult to establish in all of science, and there are very few truly secure ones in contemporary behavioral neuroscience. Locutions such as "might be involved in," which recur often in behavioral neuroscience, are an indication of the weakness of most of its linkage hypotheses.

Moreover, and perhaps most tellingly, I believe that any study of successful physical reductionism – that is, any study of the establishment of hypotheses such as that the action potential is the physical realization of the nerve impulse, or that base-pair sequences in DNA are the physical realizations of genes – will show that the reductionist program only succeeds when the variable whose physical identity is sought is already well defined and understood in its own terms, at its own level of analysis, within a theoretical framework that rests on observations at that level. As A.V. Hill's (1936) review shows, neuroscientists acquired an elaborate qualitative and quantitative understanding of the nerve impulse based on observation of its behavioral effects (muscle twitches), without recourse to non-behavioral observations. Similarly, we had an elaborate qualitative and quantitative knowledge of genes, based on the study of the inheritance of phenotypes, prior to their identification with base-pair sequences. Indeed, it is still the case, and I would argue it always will be, that the natural unit for measuring the distance between two genes is the centimorgan, not the nucleotide.

A corollary of my beliefs about the preconditions for successful physical reductionism is that neuroscientists have more to gain from economists and psychologists than the latter have to gain from the former. I see neuroeconomics as the attempt to discover the neural variables and structures that well-established economic and psychological theory tell us should be there. Insofar as economic and psychological theory do not rest on secure foundations sunk in the soil of observations native to those sciences, then behavioral neuroscientists must join in the effort to buttress them through behavioral experiments, of which this section contains many excellent examples.

Linkage hypotheses are also called bridge hypotheses, because they bridge between levels of analysis. A bridge is never stronger than its weakest buttress. If at its psychological end one finds a conceptual swamp, then there is not much prospect for strengthening it by neurobiological observation, because the weaker the bridge, the greater the question as to the relevance of those observations. It is in this spirit – to what extent can we find in the nervous system the variables and processes that economic and psychological theory say must be there? – that I believe the chapters in this section are most profitably read. An example of the influence that well-founded psychological theory has had on neuroeconomic research is the influence of the drift diffusion model of decision making (see Ratcliff and Smith, 2004) evident in many of the chapters in this section.

Neuroeconomics rests strongly on the assumption that the variables and decision-making processes that are economically consequential are found in non-human animals as well as in humans. I think this assumption will prove to be true, but it is only an assumption, and there are some observations that might give one pause before embracing it. In particular, there is little reason to think that anything resembling a market economy exists in any non-human context. A corollary is that it is unclear that prices, whose natural unit is money – a completely fungible commodity – play any role in the behavior and mental processes of any non-human animal.

Money greatly facilitates human exchange, because two different goods are often not interchangeable in many contexts. A king in need of a horse may prefer a broken-down nag to all the tea in China. Shizgal and his collaborators (for review, see Shizgal, 1997) point out that a lack of interchangeability is also a critical issue in an animal's individual economy. Animals need many different things – and the things they need are not interchangeable. No amount of sexual titillation will substitute for orange juice in the maintenance of a monkey's internal milieu, and no amount of orange juice will substitute for an encounter with a receptive female in furthering a male monkey's reproductive success. So how is it possible to get a monkey to give up orange juice in order to be allowed to look at erotic pictures, as Platt and his collaborators manage to do in the experiments described in Chapter 29?

Shizgal and his collaborators trace the origin of mechanisms that compute subjective utility to this problem inherent in the animal economy: generally speaking, an animal can only pursue one goal at a time. The different goals it might pursue are incommensurable. In order to decide which of several incommensurable goals to pursue, the factors favoring one goal or another must be reduced to a common evaluation on a unidimensional scale, the scale of subjective utility. Shizgal and his collaborators suggest that in the course of computing the subjective utility of positive behavioral outcomes, the brain computes an intensity of reinforcement signal. They identify the intensity of a reinforcement signal with the signal that mediates the rewarding effects of focal electrical stimulation of the medial forebrain bundle. This signal is carried by myelinated axons in the medial forebrain bundle, which runs through the lateral hypothalamus, connecting the ventral midbrain to the frontal lobes (Shizgal and Murray, 1989).

The intensity of reinforcement is not the same as the utility of an outcome, because the utility is affected by the rate at which reward can be obtained and the delay in obtaining it, among other things. Nonetheless, Shizgal and Murray argue that the intensity of reinforcement is like subjective utility in that it is abstracted from the sensory/perceptual qualities inherent in actual rewards, like orange juice and erotic pictures. Intensity of reinforcement is a dimension along which sexual pictures vary as one moves from those seen in textbooks on sexual hygiene to those seen in the centerfold of a men's magazine, and along which orange juice varies as one goes from an overly diluted restoration from frozen concentrate to freshly squeezed. They argue that there is a pure intensity of reinforcement signal that represents variation along this intensity dimension independent of the other attributes of an outcome, and identify this intensity of reinforcement signal with the rewarding signal in brain stimulation reward. They argue that the electrodes in the medial forebrain bundle stimulate the system that computes the utility of outcomes after the point at which the current value of those outcomes has been computed. Moreover, they show that conventional outcomes, like saline and sucrose (non-interchangeable solutions) sum their reinforcing intensities with the signal produced by the brain stimulation reward, and that reducing or eliminating hunger and salt-deprivation reduces or even reverses the intensity of reinforcement from sucrose and saline, but neither manipulation reduces the intensity of reinforcement of the brain stimulation reward.

Their hypothesis about the natural function of the neural substrate for brain stimulation reward explains one of the most salient characteristics of the brain stimulation reward phenomenon: the lack of satiation. For any natural reinforcement, enough is enough, because the reinforcement reduces the conditions that gave it value; however, for brain stimulation reward, enough is never enough. No matter how much brain stimulation reward the animal has already obtained, its appetite for more is undiminished, as would be the case for an outcome that had a high positive intensity of reinforcement no matter what the animal's physiological state.

I mention their Shizgal hypothesis in part because it is one of the more fully studied and experimentally supported hypotheses in the field of neuroeconomics. It is, I believe, the only one that rests on a demonstration that a signal artificially injected into the brain is causally effective in a manner consistent with the function ascribed to the stimulated system. The demonstration of the causal efficacy of the signals identified with a causal variable in behavior is a *sine qua non* for a secure neurobehavioral linkage hypothesis.

Another reason for describing this hypothesis and the evidence for it is that it takes us out of the cortex, where most neuroeconomic research focuses, down deep in the brain to the hypothalamus. Curiously, the dopamine neurons that carry the timed reward-expectation signals described in several chapters in this volume (see also Schultz *et al*., 1992; Fiorillo *et al*., 2003) also run in the medial forebrain bundle. More curiously still, drugs that block the D2 receptor, a postsynaptic receptor molecule in dopamine signaling pathways, block the reinforcing effect of brain stimulation reward, in relative doses predicted by their relative affinity for that receptor (Gallistel and Davis, 1983). Among the many things on the agenda of neuroeconomic researchers in coming decades must be understanding how these hypothalamic mechanisms relate to the cortical mechanisms studied by electrophysiological means in many of the chapters in this and other sections.

MATCHING

I close this introductory essay by returning, in connection with some recent work from my own laboratory, to my point about the importance of having secure psychological theories as a precondition for the neurobiological identification of the variables in those theories and the mechanisms that realize them. Among the requirements for a valid psychological theory of choice in a given paradigm is that it correctly identifies the choice as the subject represents it. A second requirement is that it correctly describes the inputs to the decision variable, and the manner in which those inputs are processed so as to yield a decision. In Chapter 31, Lee and Wang apply their neurobiologically based stochastic model of decision making to, among other things, the matching phenomenon, which is also touched on in Chapters 30 and 32, from the Newsome and Glimcher labs, respectively, both of which have studied this phenomenon from a neuroeconomic perspective. An assumption in most models of this phenomenon is that the animal's nervous system represents this as a choice between two possible options. Other assumptions are that (1) its choice is based on a running average of (2) the outcomes of its recent choices. There are reasons to doubt all three of these assumptions.

To begin with, most models of matching in the neuroeconomic literature assume discrete trials, because that is how the matching paradigm is run in this literature (for good technical reasons). However, life is not organized into trials. I believe that the real test of a model of a matching model is, first, whether it can even be applied to the free-operant matching paradigm (Herrnstein, 1961), which, as its name suggests, does not impose a trial structure on the animal's experience and responses. In this case, there are two places to forage, and either of them occasionally yields some food; the subject is free to apportion its time between the two locations however it sees fit. Subjects move back and forth at will between the two locations, apportioning their time such that the ratio of the expected durations of their stays at the two locations matches the ratio of the expected incomes from the two locations (food per unit time). Many trial-based models fail this first test, because it is unclear how to generalize them to the continuous time (no trials) situation, which must be regarded as the more natural situation.

Secondly, the paradoxical results from experiments in which subjects are first observed to match given each of two different pairs of locations (with income ratios of 1:2 and 2:4), and are then offered a choice in probe trials between a pair composed of one option from each of the two original pairs (e.g., the options with incomes in the ratio 2:2), imply that the subjects are not in fact choosing between the two options on offer. They are not making the choice they appear to be making. Rather, they are deciding when to abandon whichever option they have already chosen (Belke, 1992; Gibbon, 1995; Gallistel and Gibbon, 2000; Gallistel *et al.*, 2001).

Thirdly, and of particular interest to economists, the results from our experiments are not consistent with the assumption that matching is based on the gains from these choices – that is, on the relative profits from the behavioral investments. Critical here is the distinction between *income* and *profit* – a distinction very familiar to economists, but less so to psychologists. In free-operant matching, the income from a location is simply the amount of food obtained there per unit of time (full stop; note, *not* per unit of time spent there), while the profit is the income per unit of time spent there. Matching equates profits, not incomes. Traditionally, models have assumed that matching was the slow equilibration of profit by a hill-climbing process in which the investment ratio was adjusted until the profit ratio was 1:1 (Herrnstein and Vaughan, 1980). We have shown, however, that the process of adjustment can be extremely fast; it can go to completion within the span of a single inter-reinforcement interval at one of the locations (Gallistel *et al.*, 2001). Moreover, when computed on a visit-by-visit basis, profit is a much noisier variable (because income appears in its denominator, and its numerator, visit duration, is an exponentially distributed random variable). Thus, changes in income are apparent much sooner than are changes in profit. We showed that the adjustment to a new ratio of incomes often occurred before there was any evidence of a change in the relative profitability of visits to the two locations (Gallistel *et al.*, 2001).

The abruptness of the adjustment to changes in incomes is a challenge to models that assume a decision variable that is a running average of recent outcomes, whether those outcomes are construed as profits or incomes. Running averages cannot change from one extreme to the other between one datum and the next. An alternative model of the decision process is one in which the estimates of the incomes are based on small samples, with re-sampling only when an analysis of the datastream that looks for changes in income finds a change (Gallistel and Gibbon, 2000; Gallistel *et al.*, 2007). With such a procedure, the successive samples contain no overlapping data, so the estimate of an income ratio can change by any amount between one datum and the next.

The assumption that matching is based on income, not profit, is an extremely odd one, because it assumes that the animal does not take into account the impact of its own sampling behavior on the incomes that it observes. If it very rarely visits a high-profit location, then its income from that location will be low – lower, perhaps, than its income from a frequently visited low-profit location. This oddity has predictable consequences, which are even odder. It predicts that when the incomes from two locations are made strictly dependent on the subject's investments in them – i.e., on how long it spends at each location – then there should be an almost exclusive preference for the more profitably location (so far, so good), BUT it should be unstable: every now and then subjects should abandon the more profitable location for a prolonged stay at the less profitable location, even though in doing so they immediately reduce both their overall income and their overall profit. That prediction should make any economist sit up and take notice. I dare say most economists regard it as so unlikely that it is not worth running the experiment. We did run the experiment, and that is what we in fact observed (Gallistel *et al*., 2007). Moreover, these paradoxical reversals in preference were again extremely abrupt; they were not consistent with the assumption that preference was based on a running average of recent outcomes.

If the study of matching and its neurobiological basis is to be an important part of neuroeconomic research – and I believe it should be – then we have got to get the psychological model right. If we go looking for a running average of the profits, and if matching is not based on such a quantity, then we will look in vain. Worse yet, we may find it, and we will then erroneously assume that finding such a signal in the nervous system validates our psychological model.

References

Belke, T.W. (1992). Stimulus preference and the transitivity of preference. *Animal Learn. Behav.* 20, 401–406.

du Bois-Reymond, E. (1852). *Animal Electricity*. London: Churchill.

Edelman, G.M. (1989). *Neural Darwinism*. Oxford: Oxford University Press.

Fiorillo, C.D., Tobler, C.D., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902.

Gallistel, C.R. and Davis, A.J. (1983). Affinity for the dopamine D_2 receptor predicts neuroleptic potency in blocking the reinforcing effect of MFB stimulation. *Pharmacol. Biochem. Behav.* 19, 867–872.

Gallistel, C.R. and Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Rev.* 107, 289–344.

Gallistel, C.R., Mark, T.A., King, A.P., and Latham, P.E. (2001). The rat approximates an ideal detector of changes in rates of reward: implications for the law of effect. *J. Exp. Psychol. Animal Behav. Proc.* 27, 354–372.

Gallistel, C.R., King, A.P., Gottlieb, D. *et al*. (2007). Is matching innate? *J. Exp. Anal. Behav.* 87, 161–199.

Gibbon, J. (1995). Dynamics of time matching: arousal makes better seem worse. *Psychonomic Bull. Rev.* 2, 208–215.

Herrnstein, R.J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* 4, 267–272.

Herrnstein, RJ. and Vaughan, W.J. (1980). Melioration and behavioral allocation. In: J.E.R. Staddon (ed.), *Limits to Action: The Allocation of Individual Behavior*. New York, NY: Academic Press, pp. 143–176.

Hill, A.V. (1936). Excitation and accommodation in nerve. *Proc. R. Soc. Lond. B* 119, 305–355.

Ratcliff, R. and Smith, P.L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Rev.* 111, 333–367.

Schultz, W., Apicella, P., Scarnati, E., and Ljungberg, T. (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. *J. Neurosci.* 12, 4595–4610.

Shizgal, P. (1997). Neural basis of utility estimation. *Curr. Opin. Neurobiol.* 7, 198–208.

Shizgal, P. and Murray, B. (1989). Neuronal basis of intracranial self-stimulation. In: J.M. Liebman and S.J. Cooper (eds), *The Neuropharmacological Basis of Reward*. Oxford: Clarendon Press, pp. 106–163.

C H A P T E R

28

The Computation and Comparison of Value in Goal-directed Choice

Antonio Rangel

OUTLINE

INTRODUCTION

Neuroconomics studies the neurobiological and computational basis of value-based decision making. As the name suggests, these types of decisions involve the computation and comparison of values. This gives rise to one of the most basic questions in the field: how does the brain compute and compare values? This chapter provides an integrative review of computational and neurobiological studies addressing this question.

The chapter has several goals. It seeks to introduce economists and modelers to the neuroscience of the problem, and neuroscientists and psychologists to the relevant models. This interplay between theory and neurobiology has been very productive in many areas of neuroeconomics, and our problem is no exception[1]. The chapter also seeks to provide a framework that brings together computational and neurobiological findings by highlighting the connections between the relevant literatures in neuroscience, psychology, and economics. Finally, it seeks to motivate future work in this area by emphasizing some of the most important open questions.

Understanding how the brain computes and compares values is important for several reasons. First, in

---

[1] Caveat: in order to keep the exposition accessible to both audiences, important simplifications have been made.

order to make good decisions, organisms need to assign values to actions that are commensurate with the level of rewards that they generate, and they need to select the actions with the highest values. The extent to which they are able to do this depends on the properties of the algorithms that they use, and on the performance of the "wetware" implementing them.

Second, many models in economics (e.g., subjective experienced utility), psychology (e.g., prospect theory), and computer science (e.g., Q-learning) implicitly assume that the brain can flawlessly and instantaneously assign appropriate values to actions. Evidence from behavioral economics suggests that this is not always a correct assumption. A detailed understanding of how the brain computes and compares values should provide important insights on the strengths and limitations of existing models.

Third, evolution has endowed the brain with the capacity to make good choices in a large variety of environments, but not in all of them. In particular, organisms are susceptible to the influence of environmental variables that interfere with their ability to compute accurate values or select the best actions. Knowledge of the decision circuitry's shortcomings has implications for public policy debates, such as the regulation of marketing, and for the design of economic institutions that can help people compensate for these shortcomings.

Finally, one of the most exciting areas of application for neuroeconomics is psychiatry. Damage to various neural systems may alter the way the brain assigns and compares values. In fact, psychiatric diseases such as obsessive-compulsivedisorders, anorexia, psychopathy, and addiction can be characterized as impairments in decision making. It follows that an improved understanding of the neurobiology of value-based choice should have immediate diagnostic and therapeutic applications.

The review is organized as follows. The first section defines goal-directed choice and places the problem on the context of the neuroeconomics research agenda, while the second reviews the literature on perceptual decision making. The third section reviews the literature on goal-directed choice, before the chapter concludes with a final discussion.

PRELIMINARIES: WHAT IS GOAL-DIRECTED CHOICE AND WHAT NEEDS TO BE COMPUTED?

Goal-directed Choice

Value-based decision making is a complex process that requires, at the very least, the deployment of five basic types of computations (Rangel *et al.*, 2008). First, a representation of the decision problem needs to be constructed. This entails identifying potential courses of action (e.g., to pursue a prey item or not), as well internal states (e.g., the hunger level) and external states (e.g., the threat level). Second, a value needs to be assigned to the different actions under consideration. In order to make sound decisions, these values need to be aligned with the benefits associated with each action. Third, the valuations are compared in order to make a choice. Fourth, after implementing the decision, the brain needs to measure the desirability of the outcomes that follow. Finally, this feedback is used to update the other processes in order to improve the quality of future decisions. These are not rigid categories, but they provide a useful and reasonable decomposition of decision making into its composite parts.

Another complication in understanding value-based choice is that a sizable body of evidence has shown that the brain utilizes multiple systems to evaluate the options under consideration (for reviews, see Balleine, 2005; Dayan, 2008; Rangel *et al.*, 2008; also Chapters 12 and 24 in this volume). At least three different types of systems have been identified so far: Pavlovian, habitual, and goal-directed. These systems are deployed in different circumstances, and have different relative advantages and disadvantages.

The Pavlovian systems assign values only to a small set of "hard-wired" behaviors that are evolutionarily appropriate responses to appetitive or aversive outcomes. Typical examples include preparatory behaviors such as approaching cues that predict the delivery of food, and consummatory responses to a reward such as pecking at a food magazine. An advantage of these systems is that they are fast, which can be invaluable in many circumstances. A wide range of behaviors with important economic consequences are likely to be controlled by the Pavlovian valuation systems. Examples include overeating in the presence of food, and perhaps even the harvesting of immediate present smaller rewards at the expense of delayed non-present larger rewards (Dayan *et al.*, 2006; Dayan and Seymour, 2008).

In contrast to the Pavlovian systems, which activate only a small set of hard-wired responses, habitual systems are capable of learning to assign values to a large number of actions through a process of trial and error. These systems exhibit the following characteristics. First, such systems learn to assign values to actions that are commensurate with the expected reward that they generate as long as sufficient practice is provided and the environment is stable enough (see Sutton and Barto, 1998; also Chapter 22 in this volume). Second, the systems learn and update values slowly – thus they forecast the value of actions incorrectly whenever there is a sudden change in the rewards associated with the different actions. Third, the systems do not appear to assign values to novel actions, and thus cannot guide decision making in new situations.

Goal-directed systems are capable of computing values in novel situations and in environments with rapidly changing action–outcome contingencies. These systems assign values to actions by computing action–outcome contingencies and then evaluating the rewards associated with the different outcomes. Under ideal conditions, the value assigned to an action equals the average expected reward obtained from that action. When the correct action-contingencies are used, and the brain evaluates outcomes correctly, such systems compute an accurate forecast of the expected reward generated by the action. A key disadvantage of goal-directed systems is that they are computationally intensive, especially in comparison to the habitual and Pavlovian systems, and require a large amount of accurate information to forecast values correctly (Daw *et al.*, 2005) Furthermore, in the presence of uncertainty about the action–outcome contingencies, their predictions are noisy forecasts of the true value of the action.

This chapter makes several simplifications to make our problem tractable. First, it focuses on the valuation and action-selection stages, and ignores the representation, evaluation, and learning processes (but see Chapters 22 and 24 for a more detailed discussion of these learning processes). Second, it considers the computation and comparison of values by goal-directed systems, but not by the habitual and Pavlovian systems. Third, it focuses on a simple type of goal-directed choice in which (1) there are only two options, (2) subjects are familiar with the outcomes generated by the two actions, and (3) consumption of the outcomes is immediate. The last simplification is useful, because it allows us to ignore the issue of how the value of outcomes is modulated by time and uncertainty (for a discussion of these issues, see Rangel *et al.*, 2008, and Chapter 23 of this volume).

What Needs to be Computed?

A simple goal-directed choice involves a choice between a pair of actions – call them a and b. Subjects care about which action they take because the actions lead to different outcomes, denoted by o. The outcomes are associated with an amount of reward $v(o)$. The mapping of actions to outcomes may depend on the state of the world, denoted by s. For example, pulling the lever in a gambling machine (often called a bandit) leads to a desirable outcome if it is in the "payout" state, but to an undesirable one if it is in the "no payout" state.

In order to make good choices, the goal-directed valuation system needs to (1) assign values to actions that are commensurate with the amount of expected reward that they generate, and (2) choose the action that has been assigned the highest value. Thus, ideally the system ought to assign a value to action x equal to

$$U_x = \sum_s p(s)v(o_x(s))$$

where $p(s)$ is the probability of state s, and $o_x(s)$ is the outcome generated by the action in state s.

Note that in order to estimate this value the brain needs to calculate in real time the probability of the potential states, the action–state–outcome associations, and the outcome–reward pairings. As we will show below, making these computations takes time.

PERCEPTUAL DECISION MAKING

We begin our discussion of the computation and comparison of goal-directed values with one of the best-studied problems in behavioral neuroscience: visual-saccadic decision making. An example of this type of decision making is the now classic experimental paradigm of Newsome *et al.* (1989), referred to here as Random Dot Motion (RDM) task. As illustrated in Figure 28.1a, monkeys are shown a patch of

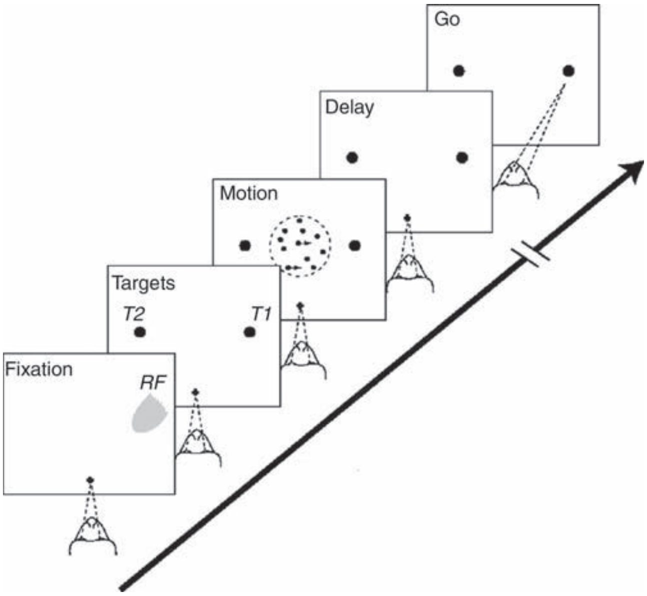


FIGURE 28.1 The Random Dot Motion task. The monkey is trained to perform a two-alternative forced choice discrimination task. He is shown a patch of randomly moving dots that contain a fraction of coherent movement to either T1 or T2, and is rewarded for correctly choosing the coherent direction. Choices are indicated by fixating, after a delay, on one of the two targets. Adapted from Shadlen and Newsome (2001), with permission.

moving dots. Most of the dots are moving randomly, but a fraction of them has a coherent direction of motion. The animal's task is to indicate which of two targets reflects the coherent direction of motion by executing an eye movement to the appropriate target. The difficulty of the task is varied across trials by changing the percentage of dots that are moving coherently. Subjects are rewarded a fixed amount for correct answers, and receive nothing for incorrect responses. (For more details on this task see Chapters 4 and 30 of this volume).

At first sight, the RDM task looks very different from value-driven choice. This is not the case. Subjects care about making the right choice because it affects the reward that they get. The value of making an eye movement equals the size of the reward for a correct response multiplied by the probability that that direction is the correct one. Subjects can estimate the action values using their senses to measure the net direction of motion in the stimulus display. This is easy when the percentage of coherent dots is high, and difficult when it is small. In either case, the task involves linking sensation to valuation and then using those values to select an action.

Note that the task is a special case of the model described above. There are two states of the world: the coherent motion goes in the direction of target 1 ($s = 1$) or it goes in the direction of target 2 ($s = 2$). The mapping of actions and states to outcome is given by:

$$o_1(s = 1) = o_2(s = \;) = \text{one unit of reward}$$

and

$$o_1(s = \;) = o_2(s = 1) = \text{zero units of reward.}$$

For simplicity, and without imposing any restrictions in the data, we can pick the units in which the reward is measured so that $v(1) = 1$ and $v(0) = 0$. It then follows that

$$U_1 = p(s = 1)v(1) + p(s = 2)v(0) = p(s = 1)$$

and

$$U_2 = p(s = 2)v(1) + p(s = 1)v(0) = p(s = 2)$$

In other words, the value of choosing one of the targets is equal to the probability that it reflects the true coherent direction of motion. It follows that the problem of computing the value of the actions is closely related to the perceptual task of estimating the the direction of motion.

Although other perceptual discrimination tasks have been studied in the literature, the RDM task is the one that is best understood from both computational and neurobiological perspectives. For this reason, the rest of this chapter focuses on the RDM task (see Romo and Salinas, 2003, and Gold and Shadlen, 2007, for reviews of alternative tasks).

Model

A large number of computational models describing how the brain computes and compares values has been proposed. These include, among others, the independent accumulators model (Luce, 1986); the random-walk and diffusion models (Stone, 1960; Ratcliff, 1978; Laming, 1979; Link, 1992; and Ratcliff *et al*., 2003); and several hybrid models (Smith, 1995; Usher and McClelland, 2001). For detailed discussions of the differences and similarities among all of them, see Smith (2000), Ratcliff and Smith (2004), Ditterich (2006), and Bogacz (2007).

All of these models share some features that are illustrated in Figure 28.2a. First, they contain processes that encode the moment-to-moment sensory evidence in support of the two potential states of the world. For example, in this figure there is a population of neurons encoding evidence in favor of the first target, and another population encoding evidence in favor of the second. Importantly, the models assume that these perceptual measurement processes are noisy. Second, there is an accumulator process associated with each of the two states that integrates the cumulative sensory evidence which it has received in its favor. Note how, as the trial progresses, the accumulators average out some of the noise. Finally, the models contain a criterion that needs to be satisfied for a choice to be made. For example, in the accumulators model, the first accumulator to reach a threshold in selected. The models listed in the previous paragraph differ in the details of the integration and comparison processes they employ to accomplish this sequence of events – see Figure 2 in Smith and Ratcliff (2004) and Figure 1 in Bogacz (2007) for useful summaries.

Here, we focus on the random-walk model (which is a version of the diffusion model in which computations are made in discrete units of time). This version of the model is attractive because it is mathematically simpler than its counterparts, and it makes predictions that can account for important aspects of the psychometric and reaction-time functions. Furthermore, a growing body of research has found signals in the brain that resemble some of the variables computed by the random-walk model (more on this below).

The random-walk model is illustrated in Figures 28.2b–c. The key variable is a measure of relative evidence: positive values of this variable denote that the

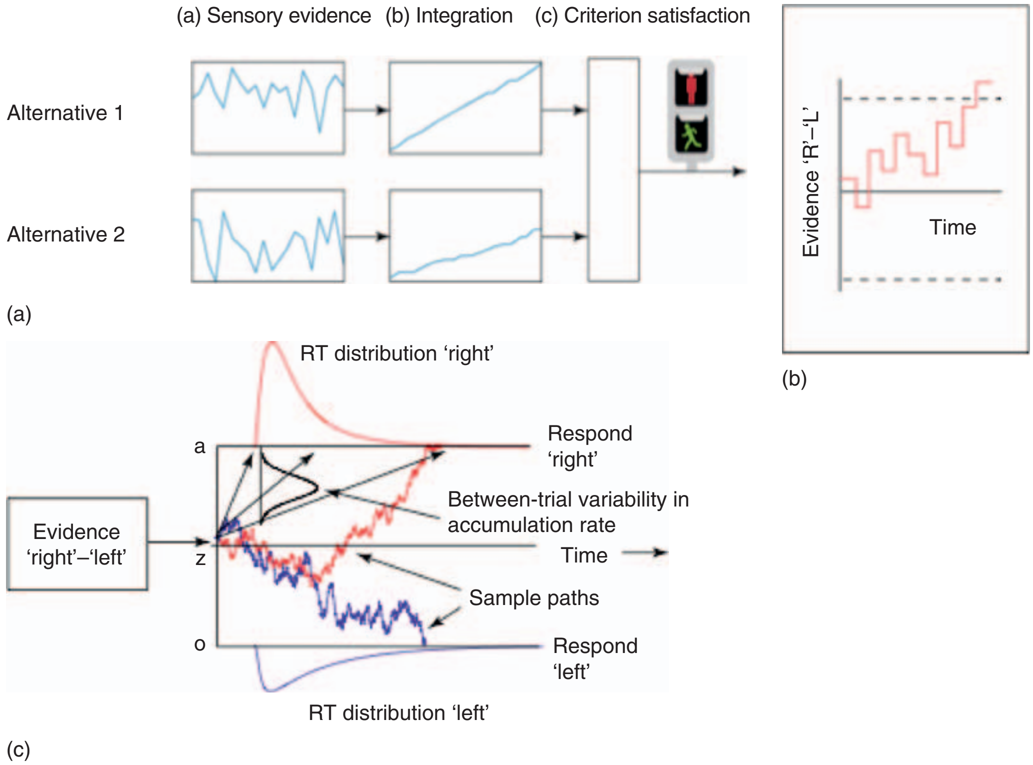


FIGURE 28.2 Models of the RDM-task. (a) Schematic illustrating the main components of the models. Reproduced from Bogacz (2007), with permission. (b) A typical run of the random-walk model. The red line represents the amount of relative evidence accumulated for the "right" target. The process starts at a middle point and stops the first time this variable crosses one of the thresholds (depicted by the bracketed horizontal lines). "Right" is chosen when it crosses the upper threshold; "left" is chosen when it crosses the lower one. Time advances in discrete steps. The size of every step is given by a Gaussian distribution with a mean that is proportional to the true direction of motion. This noise is meant to capture the variability in the perceptual processes. (c) A typical run of the continuous version of the model. Note that the red and blue lines represent two different sample paths and that, due to the variability in the model, it is possible for "left" to be chosen even if the net coherent motion goes right. Adapted from Smith and Ratcliff (2004), with permission.

estimation process favors the "right" decision, whereas negative values denote the opposite. The process starts at a middle point and stops the first time that this variable crosses one of the thresholds. Time advances in discrete steps. The size of every step is given by a Gaussian distribution with a mean that is proportional to the true direction of motion. This noise is meant to capture the variability in the sensory processes.

The model has several interesting properties. First, it predicts a logistic relationship between choice and the extent to which the dots are moving in a coherent direction. Second, it predicts that the distribution of reaction times is right-skewed, and that reaction times are longer for error trials (Figure 28.2c). Third, it also predicts that reaction times decrease with stimulus coherence and that accuracy improves. Finally, it can be shown that it implements a sequential probability ratio test (Wald and Wolfowitz, 1948), which is the optimal statistical test for the type of inference problem faced by the brain in the RDM task (see Gold and Shadlen, 2001, 2002, for details).

Although significant progress has been made in modeling this problem, many important theoretical questions remain unsolved. Perhaps the most important limitations of existing models are the fact that they only apply to the case of two alternatives, and that it is not obvious how to adapt them to the general case (for some initial attempts to resolve this issue, see McMillen and Holmes, 2005, and Bogacz and Gurney, 2007). Another problem is the proliferation of competing models in the literature. For example, in addition to those listed above, Brown and Heathcote (2005) have recently proposed a new class of "ballistic models" where the initial state for the accumulators is random but the path of integration is deterministic. An important task for the upcoming years is to use novel behavioral tests, as well neural data, to compare these models in depth. Most likely, the model that will best explain the data is yet to be developed. Finally, in order to improve the connection between the theory and the neurobiology, neuroeconomists need to develop a new generation of models that describe more closely the role of the relevant brain nuclei and neurotransmitter systems. For some important initial efforts along these lines, see Wang (2002), Brown and Heathcote (2005), Brown *et al*. (2005), Lo and Wang (2006), and Wong and Wang (2006).

Note that the models of perceptual decision-making described here are ambitious in their scope. They provide a unified theory of value computation, action selection, and reaction times. They make stark behavioral predictions about how the psychometric choice function and the reaction time distribution should be affected by the coherence of the stimulus. Most importantly, they also make predictions about which decision variables should be encoded in the brain, and about how they are computed. As we will see in the subsequent two sections, the random-walk model has been quite successful in all of these respects.

Behavioral Evidence

The behavioral predictions of the random-walk model and its variants are largely consistent with the data from the RDM task. Figure 28.3 illustrates this using data from a human version of the task (Palmer *et al*., 2005). As predicted by the model, choices are a logistic function of stimulus strength (Figure 28.3a), reaction times decrease in stimulus strength (Figure 28.3b), and response times are longer in error than in correct trials (Figure 28.3c). See Smith and Ratcliff (2004) for a further discussion of the behavioral predictions made by the different models.

Neurobiological Evidence

A central goal of neuroeconomics is to construct theories of value-based decision making that are neurobiologically grounded. Therefore, a key test of any neuroeconomic model is whether there is evidence that the computations it specifies are being carried out by the brain. Over the past 15 years, a sizable number of electrophysiology studies have probed the neural basis of the random-walk model. This section reviews the most salient findings and describes some of the open questions. For more comprehensive reviews of the neurobiology of perceptual decision making, see Platt (2002), Glimcher (2003), Romo and Salinas (2003), Schall (2005), and Gold and Shadlen (2007).

Since some readers of this chapter will be unfamiliar with the brain areas discussed below, a mini-primer of their location and known general function is provided (see Kandell *et al*., 2000, for more details). Figure 28.4 depicts two networks: the green one involves structures and connections that have been associated with

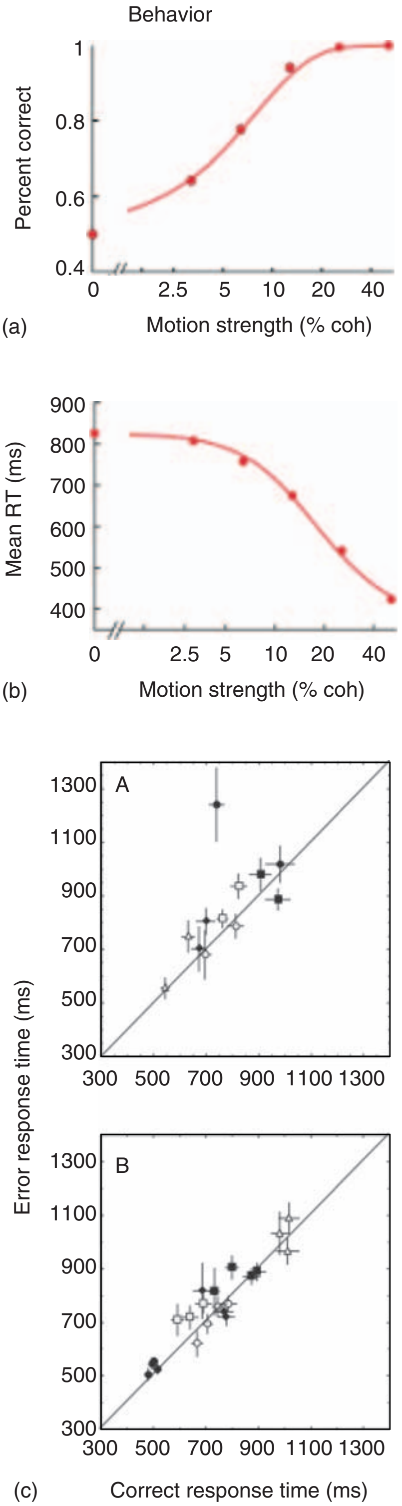


FIGURE 28.3 Behavioral predictions of the random-walk model in a human version of the RDM-task. (a) The random-walk model leads to a logistic choice function. (b) Reaction times and choice accuracy increase with stimulus coherence. Adapted from Gold and Shadlen (2007), with permission. (c) Mean response times for errors were slower than those for correct responses. Each point corresponds to a condition–observer pair. Adapted with permission from Palmer *et al*. (2005).

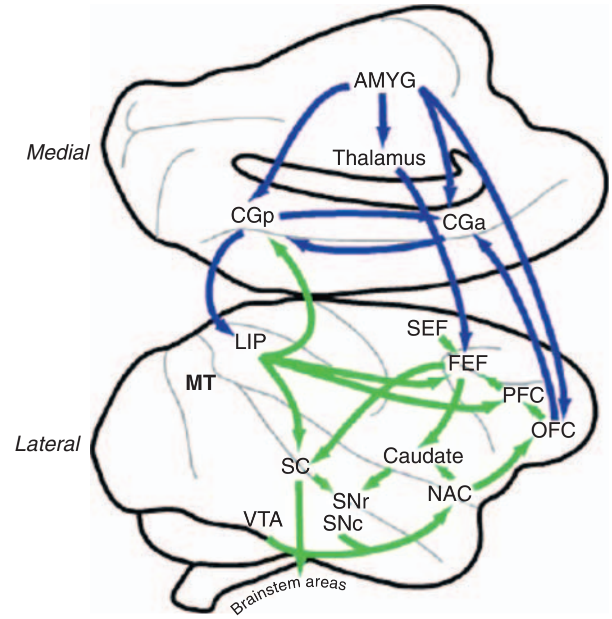


FIGURE 28.4 Network of brain areas that have been implicated in saccadic decision making. The figure depicts medial and lateral views of a monkey brain. The blue arrows denote signals that are have been implicated by previous studies in the planning and execution of eye-movements. The green arrows denote signals that have been implicated by previous studies in the computation and comparison of goal-values, as well as in other reward-related computations. AMYG, amygdala; CGp, posterior cingulated cortex; CGa, anterior cingulated cortex; LIP, lateral intraparietal area; MT, middle temporal area; SEF, supplementary eye fields; FEF, frontal eye fields; PFC, prefrontal cortex; OFC, orbitofrontal cortex; SC, superior colliculus; NAC, nucleus accumbens; SNr, substantia nigra pars reticulate; SNc, substatia nigra pars compacta; VTA, ventral tegmental area. Adapted from McCoy and Platt (2004), with permission.

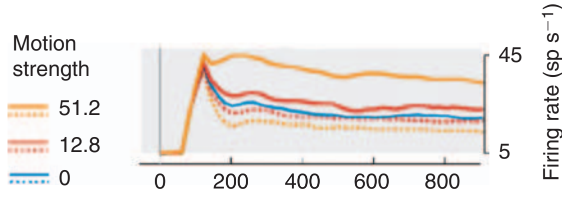


FIGURE 28.5 Activity in MT during the RDM task. The average activity in MT neurons is plotted as a function of time within the trial aligned to stimulus presentation. The panel describes the motion coherence of the stimulus. Solid and dashed curves represent trials in which the direction of motion coherence fell in and out of the neuron's response field, respectively. Note that the timecourses are fairly constant. Adapted from Gold and Shadlen (2007), with permission.

visual decision making, while the blue one involves areas that have been associated with the computation and comparison of value. It is important to emphasize that the exact roles played by the different parts of these networks are still not fully understood, and in some cases are hotly debated. Here, we focus the discussion only on the subset of these areas that have been closely associated with the computations of the random-walk model and its near relatives. Neurons in the middle temporal area (MT) encode the motion properties of objects in the visual field. The role of the lateral intraparietal area (LIP) in saccadic decision making has been the subject of past controversy: some have claimed that it is involved in spatial attention, others that it is involved primarily in the planning of saccadic eye movements. The superior colliculus (SC) is widely believed to be involved in controlling orienting behaviors such as eye movements. This area receives visual inputs from LIP, and sends motor projects to the brainstem and the spinal cord. The frontal eye-fields (FEF) are also believed to form part of the network that controls the generation of eye movements.

As we saw in Figure 28.2a, the random-walk model has three key components: a set of processes capturing the incoming sensory evidence, an accumulator measuring the amount of net evidence in favor of the alternatives, and a set of thresholds implementing the decision criterion. This provides a useful way to organize the neurobiological literature.

Consider first the sensory processes. Direction-sensitive neurons in area MT have been shown to encode information about the direction of motion in the stimulus display (Dubner and Zeki, 1971; Zeki, 1974; Britten *et al*., 1993). MT is organized around groups of neurons that respond preferentially to stimuli moving in a certain direction (Born and Bradley, 2005). Neuroscientists refer to the preferred parameters to which the group of neurons responds as the response field. Figure 28.5 illustrates the activity of a population of MT neurons with similar response fields. Activity increases with stimulus coherence when the dot movement is in the direction of the neuron's response field, and decreases with stimulus coherence otherwise. Note, as assumed by the model, that activity in the MT neurons is sustained throughout the trial. Similar results have been found for area MST (Celebrini and Newsome, 1994).

Although the previous findings are very suggestive, they do not establish a causal role for MT in the computation of goal values. After all, MT neurons could be measuring movement for purposes other than making a decision. To establish a causal role, it is necessary to use either lesion or stimulation methods. Using the later technique, Salzman *et al*. (1990, 1992) showed that microstimulation of the MT cells encoding for a particular direction of motion biased perceptual judgments in favor of that direction. Similar results were subsequently obtained for area MST (Celebrini and Newsome, 1994).

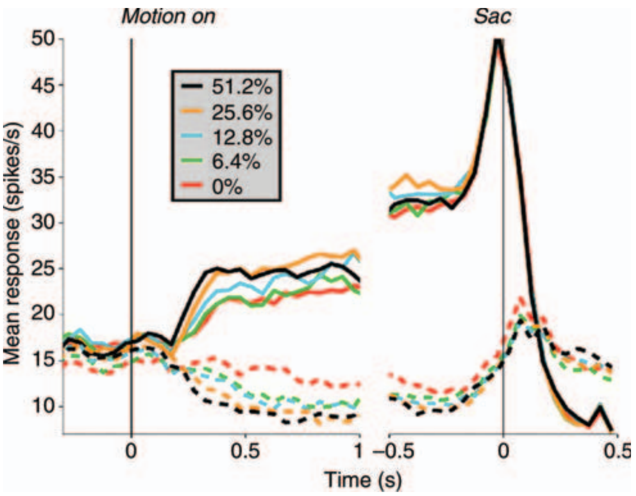


FIGURE 28.6 Activity in LIP during the RDM task. Average activity in LIP neurons is plotted as a function of time within the trial aligned to either the stimulus onset (left panel) or the time of the saccadic response (right panel). Solid and dashed curves represent trials in which the monkey judged the direction to be towards and away from the LIP response field, respectively. Note two important features of the time-courses: first, for trials in which a choice is made into the response field, firing rates are increasing on stimulus coherence. The opposite is true when the choice is out of the response field. Second, when a choice is made into the response field, all of the time-courses collapse into a single threshold about 100 ms before the initiation of the saccade. Adapted from Shadlen and Newsome (2001), with permission.

Next, consider the accumulator signal. Shadlen and Newsome (1999, 2001) found that activity in the lateral intraparietal area (LIP) resembles the accumulation of sensory signals predicted by the models. This is illustrated in Figure 28.6. Note a few things about these neurons. First, LIP neurons are also organized around response fields – in this case, LIP neurons respond to eye movements in a particular direction. Second, for trials in which a choice is made into the response field, firing rates are increasing on stimulus coherence. The opposite is true when the choice lies outside of the response field. Third, when a choice is made into the response field, all of the time-courses rise to a single threshold about 100 ms before the initiation of the saccade. A limitation of these studies is that they impose fixed response times plus a delay between stimulus presentation and response, which is inconsistent with the assumptions of the random-walk model. Follow-up experiments by Roitman and Shadlen (2002) have shown that the findings do, however, extend to the case of free and non-delayed responses.

To address the issue of causality, Hanks *et al*. (2006) microstimulated LIP neurons and found that they could bias the proportion of choices that are made towards the stimulated response field. Not only that: just as predicted by the random-walk model, stimulation generated faster reaction times when movements were made into the neurons' response field, and slower reaction times when they were made out of it. Based on this evidence, Shadlen and others have suggested that LIP neurons encode a form of the accumulator variable and that this information is used in making perceptual choices (Roitman and Shadlen, 2002; Mazurek *et al*., 2003; Huk and Shadlen, 2005; Hanks *et al*., 2006). Horwitz and Newsome (1999, 2001a, 2001b) have shown that there are cells in the superior colliculus (SC) that exhibit properties similar to those in LIP.

Finally, consider the neural implementation of the threshold. Although LIP neurons exhibit a threshold-like behavior (see Figure 28.4; Shadlen and Newsome, 2001; Roitman and Shadlen, 2002), existing experiments cannot determine whether the threshold is implemented within LIP or in other parts of the decision-making network. Lo and Wang (2006) have argued that a natural candidate for the implementation of the threshold is the SC. This midbrain area receives strong input from cortical areas, such as LIP, and contains "burst neurons" that fire just before a saccade is initiated. Lo and Wang have proposed that this cluster of neurons might be the *de facto* threshold that needs to be crossed to initiate a saccade. (For more details on this model, see Chapter 31 in this volume).

Many behavioral experiments have shown that subjects adjust their performance in response to time pressure in a way that is consistent with a decrease in the size of the thresholds (Smith and Ratcliff, 2004). This raises the question of how the threshold is modulated. Lo and Wang (2006) have proposed a model that involves a cortico-basal ganglia circuit. The SC receives inhibitory input from the substantia nigra pars reticulata (SNr), which in turn receives inhibitory input from the caudate nucleus. In the model, cortical activity increases firing rates in the caudate, which, by increasing the inhibition of the SNr, diminish the amount of activity required for the burst neurons to fire, thus decreasing the threshold. Further experiments are needed to determine the validity of this hypothesis.

Electrophysiological studies have also provided useful evidence to distinguish between the various models. Consider, for example, the experiments by Ditterich *et al*. (2003), which showed that microstimulation of MT speeds up decisions in the preferred direction of the neurons but slows down saccades in the opposite direction. These findings are consistent with the existence of a common accumulator, as in the random-walk model, but not with the existence of independent accumulators, as in the race-to-barrier model. A similar implication follows from the microstimulation experiments in LIP by Hanks *et al*. (2006).

Even though saccadic decision making is one of the better understood problems in neuroeconomics and behavioral neuroscience, much work remains to be done. First, how do neural circuits integrate the sensory information to generate the accumulator signal that has been found in LIP? Do neural ensembles in this area compute it, or does it arise from the interaction of a larger circuit that remains to be characterized? Furthermore, what is the functional difference between the accumulator signals found in LIP and those found in the SC and the frontal eye-fields? Second, it has been shown that activity in LIP is modulated by prior beliefs and expected rewards (Platt and Glimcher, 1999). These variables change the equation for the goal values described in the model above. An important open problem is to understand how these types of variables are computed and how they become incorporated into the accumulator signal found in LIP. Third, another interesting puzzle has to do with the existence of multiple accumulator signals in LIP, one for each favored direction. How do all of these signals become combined into a single decision? Existing models are silent on this, because they assume that there is a single accumulator signal. Fourth, what is the selection criterion used by the brain, and how is it implemented? The findings reviewed above are consistent with the existence of thresholds. However, it is also possible that the observed thresholds are epiphenomena and that the underlying selection criterion is substantially different. In addition, how do executive control processes modulate the selection criteria? Finally, does the random-walk model capture all of the computations required to make perceptual decisions, or are some essential processes missing?

SIMPLE GOAL-DIRECTED CHOICE

Now consider the case of simple goal-directed choice. A typical example is shown in Figure 28.7. Subjects see a pair of pictures of familiar snacks and have to decide which one to eat at the end of the experiment. They may signal their choice by making an eye movement to one of the items, through a voice command, or by pressing one of two buttons. As in the previous section, the discussion is simplified by focusing on this particular type of goal-directed choice.

The simple model of valuation derived above highlights some of the key computations that need to be made. In this case there are two actions, "left" and "right". There is only one state of the world given by the nutritional demands of the subject at the time of the decision[2]. Let o_{left} denote the snack associated with the "left" action and o_{right} that associated with the "right" choice. The value of the two actions is given by

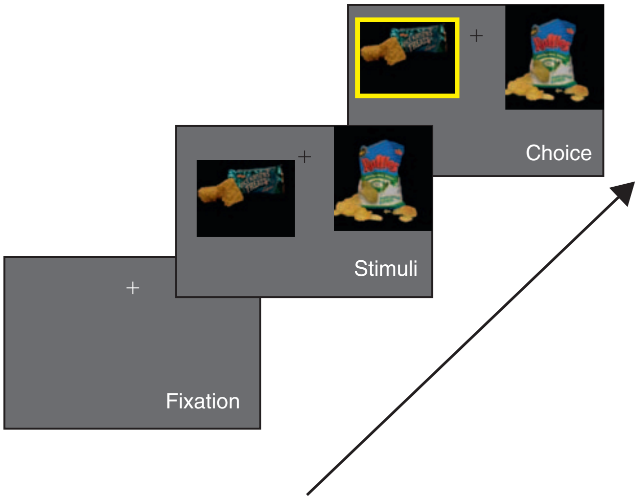


FIGURE 28.7 An example of simple goal-directed choice. After an initial fixation cross, pictures of two snacks are displayed. Subjects are then free to look at the items until they are ready to make a choice by pressing one of two keys. The pattern of fixations during the deliberation process is recorded with an eye-tracker. Adapted from Krajbich *et al.* (2007), with permission.

$$U_{\text{left}} = v(o_{\text{left}}) \text{ and } U_{\text{right}} = v(o_{\text{right}}).$$

Note some of the key differences from the RDM task. The value of an action is simply the reward obtained by eating the food to which it leads. Unlike the case of perceptual choice, the values do not depend on experimenter-controlled states that need to be estimated using sensory information. However, since the snacks associated with each action change from trial to trial, the brain needs to (1) identify the outcome associated with each action and (2) estimate its reward value. This is in contrast to the case of perceptual choice, in which the outcomes associated with the actions were constant but there was uncertainty about their probability. Despite the differences, in both cases the choice problem boils down to a computation and comparison of value signals.

Note that since no two consumption situations are identical, every time an action is evaluated the brain needs to characterize anew the outcome associated with it. For example, are the current snacks safe? What exact caloric value do they have? Since not all of the characteristics are observable before consuming the good (e.g., taste), there is an inescapable amount of uncertainty associated with every choice – for example, the trustworthiness of the experimenter and recent news about the safety of the food supply might affect

---

[2] For simplification, we omit the state from the notation below.

evaluations of the outcome. Understanding how the brain deals with this type of uncertainty is an important open question beyond the scope of this review. Here, the model and discussion are simplified by assuming that this type of uncertainty is not present, which is a reasonable approximation for the experiments discussed below.

A more important difference from the RDM-task has to do with the role of attention. Although animals can make goal-directed choices among items while maintaining a central fixation (Padoa-Schioppa and Assad, 2006), subjects normally pay sequential attention to two stimuli during the process of choice – for example, they might fixate first on the left item, observe it for a few hundred milliseconds, and then switch to the other one. This introduces a role for attention, and makes the problem of simple goal-directed choice significantly more complicated.

Theory

This section reviews the work of Krajbich and colleagues (2007), which has extended the random-walk model of perceptual decision making to the case of attention-driven choice. The idea behind the model is depicted in Figure 28.8a. The brain computes a relative value variable that is analogous to the accumulator of the random-walk model. This variable measures the relative value of "right" action. A decision is made when the decision value crosses one of the two thresholds (upper for "left" and lower for "right").

The key innovation of the model has to do with the role of attention. The relative value variable follows a random-walk process with Gaussian noise in which the mean rate of change is determined by the item that the subject is looking at. In particular, the variable changes from one instant to another according to the formula

$$r(t) = (t - 1) + \lambda(v(o_{\text{target}}) - \theta v(o_{\text{non-target}})) + \varepsilon(t)$$

where $\varepsilon(t)$ is white Gaussian noise, $v(o_{\text{target}})$ is the value of the item currently being attended, $v(o_{\text{non-target}})$ is the value of the other item, θ is a positive constant between 0 and 1 measuring the advantage given to the item being attended in the value construction process, and λ is a positive constant that scales the value functions (see Krajbich *et al.*, 2007, for further details).

The model of visual attention is extremely simple. The identity of the first fixation is random, and independent of the value of the two items. Fixation lengths are drawn from an equal distribution, and are also assumed to be unaffected by the value of the items. Note that the subject fixates on an item and then alternates fixations between both of them until a barrier is reached.

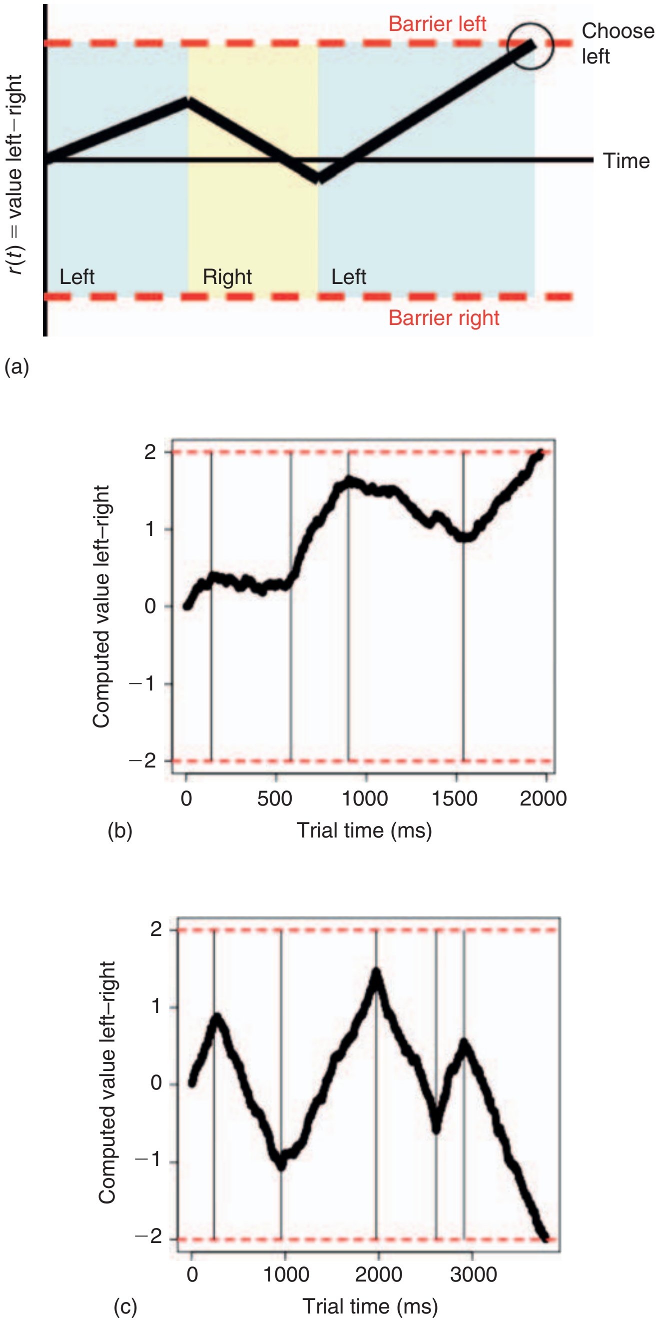


FIGURE 28.8 Model of simple goal-directed choice. (a) A typical run of the model. The thick black line represents the relative value of the "right" minus the "left" actions. The white and yellow vertical bands represent the time of the fixation. In the example depicted here, the fixations are left, right, left. The process starts at a relative value of zero and accumulates following a random-walk model that depends on the direction of the fixation. In particular, subject to the qualifications described in the text, the process approaches the left barrier while the subject is looking at the left item, and *vice versa*. A decision is made when the relative decision value signal crosses a barrier. Not shown are the random fluctuations associated with the random walk. (b) A simulated run of the model using parameters estimated from real data for a trial in which the "left" items is more valuable. (c) Another simulated run of for a trial in which both items have equal value. See Krajbich *et al.* (2007) for details.

Figures 28.8b and 28.8.c depict simulated runs with parameters estimated from real experimental data. The left panel depicts a case in which $v(o_{\text{left}}) > v(o_{\text{right}})$. Here, the signal climbs on average toward the left barrier and that choice is made with higher probability. The right panel depicts a case in which $v(o_{\text{left}}) = v(o_{\text{right}})$. In this case, the signal can fluctuate widely between the two barriers before a choice is made, and both options are selected with equal probability.

The model's predictions regarding the psychometric choice function and reaction-time profiles are similar to those of the random-walk model: choice probabilities are a logistic function of the underlying difficulty of the choices (as measured by $v(o_{\text{left}}) - v(o_{\text{right}})$), reaction times decrease with the difficulty of the choice, and reaction times are larger for error trials (in which the item with the lowest value is chosen) than in correct trials.

The model also makes predictions about the connection between visual attention and choice. For example, it predicts that last fixations are shorter (because they are interrupted when a barrier is reached), that the last fixation is to the chosen item as long as it is sufficiently better than the item seen in the next-to-last trial, and that there is a choice bias in favor of items seen first and last.

Although this model is able to account for several important features of the data, as we will see below, it leaves unanswered several important questions about goal-directed choice. First, and most importantly, is the issue of why it takes time to construct the relative value signal. In the case of perceptual decision making, there is a perceptual state that needs to be measured. However, what are the analogous processes here? One possibility is that the value signal might be computed by a process that involves observing the characteristics of the outcomes, retrieving from memory reward values associated with those characteristics, and finally integrating them. Each of these more fundamental steps might play the analogous role to sensory evidence. Second, the model assumes that the fixations are random, but this is probably not the case – for example, the first fixation might be driven by low-level visual features of the stimuli. Given the central role that visual attention plays in this class of models, it is essential to improve our understanding of the processes guiding visual attention. Third, since most choice situations involve more than two alternatives, the theory needs to be extended in this direction. For a preliminary attempt, see Pulst-Korenhberg and Rangel (2008).

The theory of attention-driven goal-directed choice presented here is not the only one that has been proposed. In particular, the model described here is closely related to decision-field theory (Roe *et al.*, 2001; Busemeyer and Towsend, 1993; Busemeyer and Johnson, 2004). The main difference between this model and decision-field theory is that whereas the former is an extension of the random-walk model, the later is more closely related to the hybrid connectionist models of McClelland and Rumelhart (1981) and Usher and McClelland (2001).

Behavioral and Eye-tracking Evidence

Krajbich and colleagues (2007) tested the behavioral implications of the model using eye-tracking as well as behavioral responses and reaction times. Figure 28.9 illustrates some of their results. The first panel shows that reaction times increase with the difficulty of the choice, as measured by the difference in an independently obtained measure of the value of the items. The second panel shows that last fixations are shorter. The third panel shows that the last fixation tends to be to the chosen item as long as the last-seen item is sufficiently better than the next-to-last item. Finally, the last panel shows that there is a duration fixation bias: the probability of choosing the left item increases with the relative amount of time for which it was seen during the choice process.

An interesting implication of the model is that it should be possible to manipulate choices by *exogenously* manipulating visual attention. Armel and colleagues (2008) have tested this prediction using a version of the experiment in which the fixation durations are controlled by the experimenter. Consistent with the predictions of the model, they show that longer exposure times increase the probability that appetitive items be chosen, but decrease the probability of choice for aversive items.

Neural Basis for the Model's Components

As in the case of perceptual decision making, a key test of the model of is the presence of neurobiological signals that resemble the computations that it makes. Unfortunately, much is less is known about the neural basis of this type of choice.

Several studies have found that activity in the orbitofrontal cortex (OFC) is correlated with behavioral measures of the goal-value assigned to objects. For example, in an fMRI study depicted in Figure 28.10, Plassmann *et al.* (2007) showed pictures of snacks to hungry subjects, who then placed bids for the right to eat them at the end of the experiment. The size of the bids was positively correlated with BOLD activity in the OFC. For related fMRI findings, see Erk *et al.* (2002), Arana *et al.* (2003), Paulus and Frank (2003),

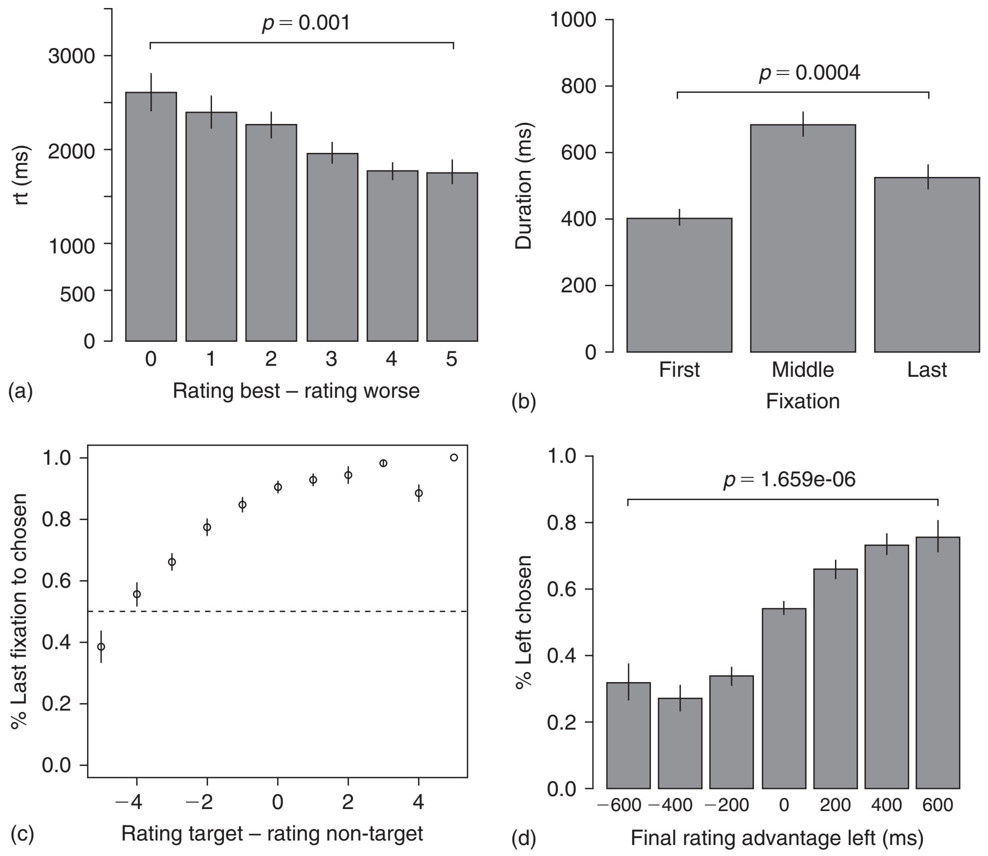


FIGURE 28.9 Eye-tracking and behavioral predictions of the goal-directed choice model. (a) Reaction times increase with the difficulty of the choice, as measured by the difference in an independently obtained measure of value. (b) Last fixations are shorter. (c) The last fixation tends to be to the chosen item, unless the last seen item is sufficiently worse than that seen next-to-last. (d) There is a duration fixation bias: the probability of choosing the left item increases with the relative amount of time for which it was seen during the choice process. See Krajbich *et al.* (2007) for more details, predictions, and analyses.

Valentin *et al.* (2007), and Hare *et al.* (2008). In a closely related monkey electrophysiology study of binary choice, Padoa-Schioppa and Assad (2006) showed that a population of neurons in OFC encoded for the subjective value of the two outcomes independently of the action that had to be taken to obtain them. Also related, Fellows and Farah (2007) showed that patients with lesions to the OFC were incapable of making consistent hypothetical food choices. Together, these findings suggest that activity in the medial OFC encodes for the value of the outcomes associated with the actions and, based on the study of Fellows and Farah, that these computations are essential for making consistent goal-directed choices.

Activity correlated with the value of outcomes has also been recorded in the dorsolateral prefrontal cortex (Leon and Shadlen, 1999; Wallis and Miller, 2003; Hare *et al.*, 2007; Plassmann *et al.*, 2007). Camus and colleagues (2008) have recently shown that disruption of activity in the right dorsolateral prefrontal cortex using rTMS interferes with subjects' ability to compute the goal-values of items. The relationship between the value computations in DLPFC and in OFC is an important open question.

A crucial question for the models described above is whether we can find a neural signal resembling the relative value variable. Pioneering studies in this area include those by Platt and Glimcher (1999), Dorris and Glimcher (2004), and Sugrue *et al.* (2004), which have recorded neurons in monkey LIP during a goal-directed choice task. Figure 28.11a describes the experimental set-up of Sugrue *et al.* (2004), which is a little different from the simple goal-directed choices that we have discussed so far. Monkeys make choices between two saccade targets (red and green), and the location of the targets is randomized in every trial. The likelihood of getting a juice reward after selecting a target depends on the history of previous actions. Choices are indicated by making a saccade to either the left or right targets. Note that this task closely resembles

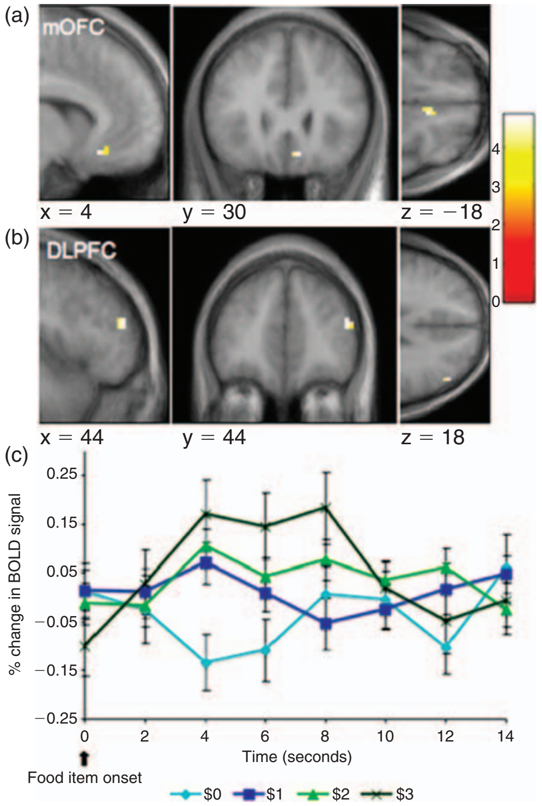


FIGURE 28.10 fMRI study of goal-directed valuation in orbitofrontal cortex. Hungry subjects were shown pictures of snacks and placed bids for the right to eat them at the end of the experiment. (a, b) The size of the bids, which are a measure of the goal values computed by the brain, were positively correlated with BOLD activity in the orbitofrontal cortex and right dorsolateral prefrontal cortex. (c) The average time-course of the BOLD signal is higher in reponse to items with larger goal values. Adapted from Plassmann *et al*. (2007), with permission.

the DRM task, except that the value of the targets now depends on the history of actions, and not on the state of a visual stimulus. Figure 28.11b depicts activity in LIP during the course of a trial, and several aspects of the result are worth highlighting. First, there are strong parallels with the case of perceptual decision making, which is perhaps not surprising, given that both are instances of goal-directed choice. Second, activity in LIP is increasing with the relative value of the value of the target in the response field, and it follows a time-course that is remarkably similar to the one in Figure 28.6 – the key difference being that activity is now sorted by the value of the target instead of the strength of the stimulus. Third, activity in LIP seems to reach a threshold just before a movement into the neuron's response field, also just as before. Although so far the evidence is only correlational, the parallel with the RDM task suggests that LIP is encoding something akin to the relative value signal (Dorris and Glimcher, 2004). For more details on these experiments, see Chapter 31 of this volume.

Note that since the monkey maintains a central fixation during the evaluation of the targets, it is unclear whether attention plays a role in the task. One possibility is that the monkey is using covert attention in a way that is not captured by the aggregated data analysis; another is that information about the two values is being received and processed in parallel.

The excitement of these findings not withstanding, there are many open questions regarding the neurobiology of goal-directed choice. First, what is the role of the signal encoded in LIP? Second, as discussed above, an integrator-like signal was found in the SC and FEF in the RDM task. Are similar signals present in goal-directed choice? Since these areas seem to be involved in the preparation and execution of eye movements, it is natural to speculate that this will be the case when the goal-directed choices are made through a saccade, but not when they require other types of motor output (e.g., a vocal command or a hand movement). Third, which brain regions perform computations that are specific to each type of motor output, and which perform computations that apply regardless of how the action is executed? Although much research remains to be done, the existing evidence suggests the following conjecture. Areas that are primarily involved in sensory-specific processing, such as MT in the case of vision, should play a central role in perceptual choice, but not in general goal-directed choice. Areas that play a central role in the preparation and execution of particular motor outputs, such as LIP, FEF, and SC in the case of eye movements, should play a central role in both types of choices as long as the decision is expressed through that motor system. Finally, areas of the prefrontal cortex such as the DLPFC and the OFC, as well as parts of the basal ganglia not discussed here, should play a role in both types of decisions regardless of the system used to signal the choice. (For a detailed review of the role of the basal ganglia, see Chapters 25 and 32 of this volume). Fourth, the models presented in this section assume that choices are made over actions, not over outcomes. Is this a justified assumption? In the absence of further evidence, an equally plausible possibility is for the choices to be made over outcomes, and then for the necessary motor commands to be executed. Another possibility is for the brain to make choices simultaneously over outcomes and actions, and then somehow to coordinate the result. Finally, how are considerations such as risk, uncertainty, time, and social effects incorporated into the machinery of goal-directed choice? (For more on this point, see Chapter 29 of this volume.)

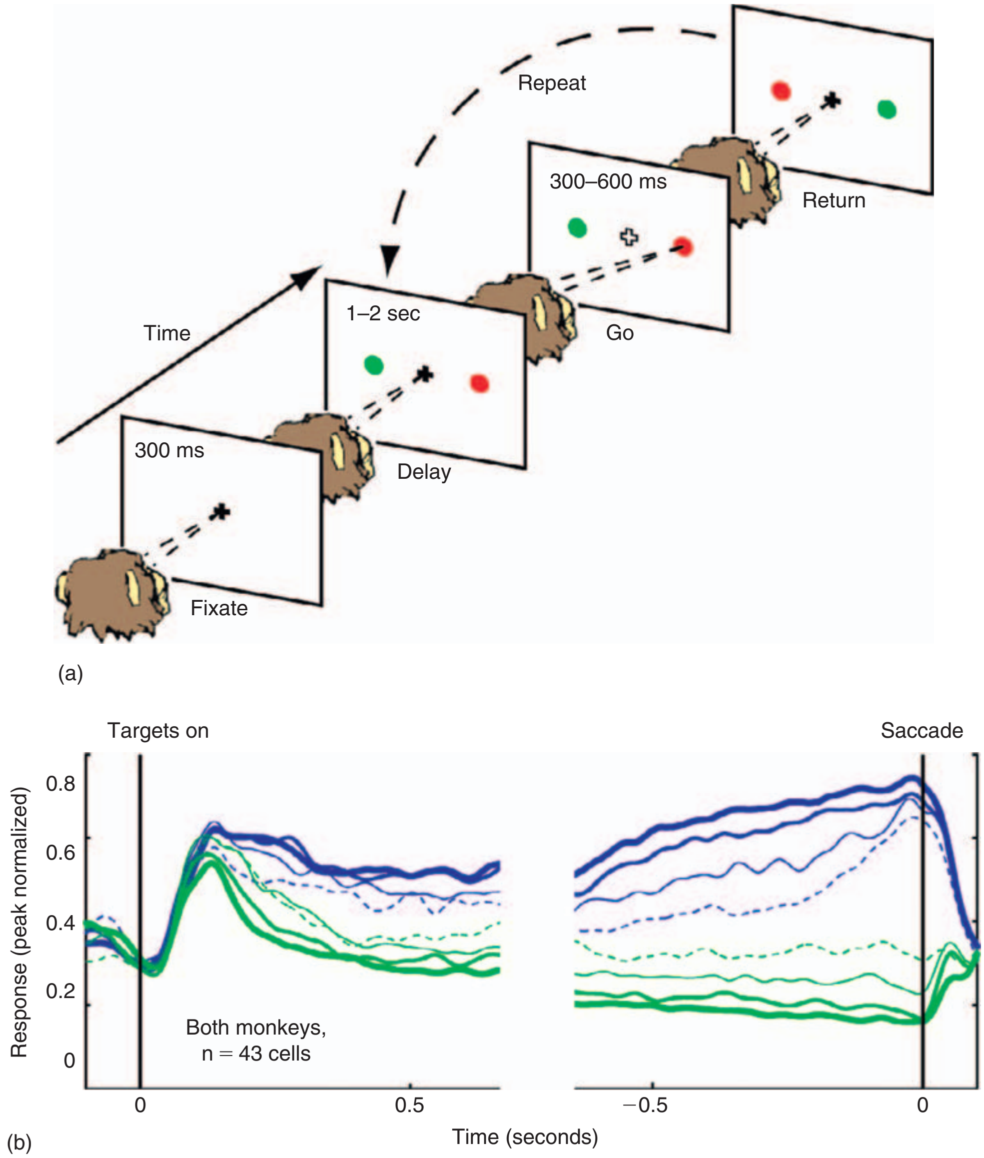


FIGURE 28.11 Activity in LIP during a goal-directed choice task. (a) The task. To begin a run of trials, the animal must fixate on the central cross. Saccade targets (red and green) then appear at randomized locations while the animal maintains fixation. Dimming of the fixation cross then signals that a saccadic response can be made. After a period of maintained fixation, the targets appear again at a random location and the trial starts again. The reward from choosing a target depends on the history of previous actions (see Sugrue *et al.*, 2004, for details). (b) Average activity in LIP neurons is plotted as a function of time and is aligned to either the appearance of the targets at the beginning of the trial (left panel) or to the time of the saccadic response (right panel). Blue curves denote trials in which the monkey made a choice into the neuron's response field. Green curves denote trials in which the choice was outside of the response field. In both cases, thicker lines indicate that the chosen target has a higher associated expected reward. Note that these time-courses closely resemble those in Figure 28.3; the only difference is that now LIP is discriminating based on the expected value of the chosen targets instead of stimulus coherence.

FINAL REMARKS

A central goal of neuroeconomics is to construct a theory of value-based choice that is grounded in neurobiology. The problem of the construction and comparison of values, especially for the special case of perceptual decision making, is perhaps the most advanced example of this research agenda. As has been shown in this review, we now have models describing the computations that the brain makes in order to make simple perceptual and goal-directed choices. The models make adequate behavioral predictions about choices, reaction times, and fixation patterns in a wide class of experimental settings. Furthermore, electrophysiology, fMRI, rTMS, lesion, and microstimulation studies are beginning to demonstrate that some of the key computations made in the models are instantiated in the brain. Thus, although much remains to be done, there is reason to hope that rapid progress will be made on these issues over the next decade.

Acknowledgments

Financial support from the NSF (SES-0134618) and the Gordon and Betty Moore Foundation is gratefully acknowledged. I thank Paul Glimcher, Colin Camerer, and Todd Hare for their useful comments.

References

Arana, F.S., Parkinson, J.A., Hinton, E. *et al*. (2003). Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection. *J. Neurosci.* 23, 9632–9638.

Armel, C., Beaumel, A., and Rangel, A. (2008). The impact of computation time on decision values. *Am. Econ. Rev.*, (forthcoming).

Balleine, B.W. (2005). Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. *Physiol. Behav.* 86, 717–730.

Bogacz, R. (2007). Optimal decision-making theories: linking neurobiology with behaviour. *Trends Cogn. Sci.* 11, 118–125.

Bogacz, R. and Gurney, K. (2007). The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural. Comput.* 19, 442–477.

Born, R.T. and Bradley, D.C. (2005). Structure and function of visual area MT. *Annu. Rev. Neurosci.* 28, 157–189.

Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.* 10, 1157–1169.

Brown, E., Gao, J., Holmes, P. *et al*. (2005). Simple neural networks that optimize decisions. *Intl J. Bifurc. Chaos* 15, 803–826.

Brown, S. and Heathcote, A. (2005). A ballistic model of choice response time. *Psychological Rev.* 112, 117–128.

Busemeyer, J.R. and Johnson, J.G. (2004). Computational models of decision making. In: D. Koehler and N. Narvey (eds), *Handbook of Judgment and Decision Making*. New York, NY: Blackwell Publishing Co., pp. 133–154.

Busemeyer, J.R. and Townsend, J.T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Rev.* 100, 432–459.

Camus, M., Shimojo, S., Camerer, C. *et al.* (2008). rTMS of right dorsolateral prefrontal cortex disrupts the computation of goal-directed values. Working Paper, Caltech.

Celebrini, S. and Newsome, W.T. (1994). Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. *J. Neurosci.* 14, 4109–4124.

Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711.

Dayan, P. (2008). The role of value systems in decision making. In: C. Engel and W. Singer (eds), *Better Than Conscious? Implications for Performance and Institutional Analysis*. Cambridge, MA: MIT Press, (in press).

Dayan, P., Niv, Y., Seymour, B., and Dayan, P. (2006). The misbehavior of value and the discipline of the will. *Neural Networks* 19, 1153–1160.

Ditterich, J. (2006). Stochastic models of decisions about motion direction: behavior and physiology. *Neural Networks* 19, 981–1012.

Ditterich, J., Mazurek, M.E., and Shadlen, M.N. (2003). Microstimulation of visual cortex affects the speed of perceptual decisions. *Nat. Neurosci.* 6, 891–898.

Dorris, M.C. and Glimcher, P.W. (2004). Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44, 365–378.

Dubner, R. and Zeki, S.M. (1971). Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Res.* 35, 528–532.

Erk, S., Spitzer, M., Wunderlich, A. *et al*. (2002). Cultural objects modulate reward circuitry. *NeuroReport* 13, 2499–2503.

Fellows, L.K. and Farah, M.J. (2007). The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment *per se*? *Cerebral Cortex* 17(11), 2669–2674.

Glimcher, P.W. (2003). The neurobiology of visual-saccadic decision making. *Annu. Rev. Neurosci.* 26, 133–179.

Gold, J.I. and Shadlen, M.N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* 5, 10–16.

Gold, J.I. and Shadlen, M.N. (2002). Banburisms and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* 36, 299–308.

Gold, J.I. and Shadlen, M.N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 374–535.

Hanks, T.D., Ditterich, J., and Shadlen, M.N. (2006). Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nat. Neurosci.* 9, 682–689.

Hare, T., O'Doherty, J., Camerer, C. *et al*. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.*, (forthcoming).

Horwitz, G.D. and Newsome, W.T. (1999). Separate signals for target selection and movement specification in the superior colliculus. *Science* 284, 1158–1161.

Horwitz, G.D. and Newsome, W.T. (2001a). Target selection for saccadic eye movements: direction-selective visual responses in the superior colliculus. *J. Neurophysiol.* 86, 2527–2542.

Horwitz, G.D. and Newsome, W.T. (2001b). Target selection for saccadic eye movements: prelude activity in the superior colliculus during a direction-discrimination task. *J. Neurophysiol.* 86, 2543–2558.

Huk, A.C. and Shadlen, M.N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J. Neurosci.* 25, 10420–10436.

Kandell, E.R., Schwartz, J., and Jessell, T. (2002). *Principles of Neural Science*. New York, NY: McGraw-Hill.

Krajbich, I., Armel, C., and Rangel, A. (2007). Visual attention drives the computation of value in goal–directed choice. Working Paper, Caltech.

Laming, D. (1979). A critical comparison of two random-walk models for choice reaction time. *Acta Psychologica* 43, 431–453.

Leon, M.I. and Shadlen, M.N. (1999). Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* 24, 415–425.

Link, S.W. (1992). *The Wave Theory of Difference and Similarity*. Hillsdale, NJ: Lawrence Erlbaum.

Lo, C.C. and Wang, X.J. (2006). Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat. Neurosci.* 9, 956–963.

Luce, R.D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford: Oxford University Press.

Mazurek, M.E., Roitman, J.D., Ditterich, J., and Shadlen, M.N. (2003). A role for neural integrators in perceptual decision making. *Cerebral Cortex* 13, 1257–1269.

McClelland, J. and Rumelhart, D. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Rev.* 88, 375–407.

McCoy, A. and Platt, M.L. (2005). Expectations and outcomes: decision-making in the primate brain. *J. Comp. Physiol. A. Neuroethol. Sens. Neural Behav. Physiol.* 191, 201–211.

McMillen, T. and Holmes, P. (2005). The dynamics of choice among multiple alternatives. *J. Math. Psychol.* 50, 30–57.

Newsome, W.T., Britten, K.H., and Movshon, J.A. (1989). Neuronal correlates of a perceptual decision. *Nature* 341, 52–54.

Padoa-Schioppa, C. and Assad, J.A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.

Palmer, J., Huk, A.C., and Shadlen, M.N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *J. Vision* 5, 376–404.

Paulus, M.P. and Frank, L.R. (2003). Ventromedial prefrontal cortex activation is critical for preference judgments. *NeuroReport* 14, 1311–1315.

Plassmann, H., O'Doherty, J., and Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27, 9984–9988.

Platt, M.L. (2002). Neural correlates of decisions. *Curr. Opin. Neurobiol.* 12, 141–148.

Platt, M.L. and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.

Pulst-Korenhberg, J. and Rangel, A. (2008). The role of visual attention in trinary economic choice. Working Paper, Caltech.

Rangel, A., Camerer, C., and Montague, R. (2008). A framework for studying the neurobiology of value-based decision making. Forthcoming in *Nature Reviews Neuroscience*.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Rev.* 85, 59–108.

Ratcliff, R., Cherian, A., and Segraves, M. (2003). A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two-choice decisions. *J. Neurophysiol.* 90, 1392–1407.

Ratcliff, R. and Smith, P. (2004). A comparison of sequential sampling modles for two-choice reaction time. *Psychological Rev.* 111, 333–367.

Roe, R.M., Busemeyer, J., and Townsend, J.T. (2001). Multialternative decision field theory: a dynamic connectionist model of decision making. *Psychological Rev.* 108, 370–392.

Roitman, J.D. and Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489.

Romo, R. and Salinas, E. (2003). Flutter discrimination: neural codes, perception, memory and decision making. *Nat. Rev. Neurosci.* 4, 203–218.

Salzman, C.D., Britten, K.H., and Newsome, W.T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. *Nature* 346, 174–177.

Salzman, C.D., Murasugi, C.M., Britten, K.H., and Newsome, W.T. (1992). Microstimulation in visual area MT: effects on direction discrimination performance. *J. Neurosci.* 12, 2331–2355.

Schall, J.D. (2005). Decision making. *Curr. Biol.* 15, R9–11.

Shadlen, M.N. and Newsome, W.T. (1996). Motion perception: seeing and deciding. *Proc. Natl Acad. Sci. USA.* 93, 628–633.

Shadlen, M.N. and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* 86, 1916–1936.

Smith, P. (1995). Psychophysically-principled models of visual simple reaction time. *Psychological Rev.* 102, 567–593.

Smith, P. (2000). Stochastic dynamic models of response time and accuracy: a foundational primer. *J. Math. Psychol.* 44, 408–463.

Smith, P.L. and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161–168.

Stone, M. (1960). Models for choice-reaction time. *Psychometrika* 25, 251–260.

Sugrue, L.P., Corrado, G., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.

Sutton, R.S. and Barto, A.G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Usher, M. and McClelland, J. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological Rev.* 108, 550–592.

Valentin, V.V., Dickinson, A., and O'Doherty, J. (2007). Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* 27, 4019–4026.

Wald, A. and Wolfowitz, J. (1948). Optimum character of the sequential probability ration test. *Ann. Math. Stats* 19, 326–339.

Wallis, J.D. and Miller, E.K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur. J. Neurosci.* 18, 2069–2081.

Wang, X.J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36, 955–968.

Wong, K.F. and Wang, X.J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *J. Neurosci.* 26, 1314–1328.

Zeki, S.M. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J. Physiology,* 236, 549–573.

CHAPTER

29

Neuronal Representations of Value

Michael Platt and Camillo Padoa-Schioppa

OUTLINE

INTRODUCTION

Value and Decision Making in Evolutionary Perspective

Humans continuously engage in behaviors that entail a choice. For example, economic choice – the choice between different goods – takes place when people buy groceries, select a pension plan, or choose a spouse. Choices also take place in the perceptual and motor domains. For example, in the presence of ambiguous sensory stimuli, a particular percept can be thought of as the result of a choice process (Gold and Shadlen, 2001). Moreover, perceptual attention – the selection of one particular stimulus to which processing resources are preferentially allocated – can also be described as the result of a choice process. In the motor domain, enacting an abstract plan (e.g., "pick an apple") entails selecting one particular action (e.g., "reach with the left hand") out of many suitable actions, which implies a choice.

The brain systems that generate such choices originally evolved to promote behaviors that enhance biological fitness, such as acquiring food and shelter, attracting mates, avoiding predators, and prevailing over competitors. Thus, the nervous system comprises a suite of morphological and behavioral adaptations for surmounting specific environmental and social challenges. Both theoretical and empirical studies in animals support the idea that, to the limits of physiological and cognitive constraints, behavioral choices serve to optimize evolutionary fitness and thus can be viewed as economical. Accordingly, brains appear to be exquisitely specialized to attend to key features of the environment, determine the predictive value of these features, and then use this information to compute the optimal behavioral choice. Economic concepts such as "value" and "optimality" thus provide a broad framework to describe different kinds of choice behavior.

Rewards can be considered proximate goals that, when acquired, tend to enhance survival and reproductive success. Similarly, avoiding punishment is a proximate goal that ultimately serves to enhance the long-term likelihood of survival and reproduction. These definitions extend the traditional psychological and neurobiological notions of reward and punishment, which are typically defined by the quality of eliciting approach and avoidance. Although still in the early stages, the convergence of concepts and experimental approaches from economics, psychology, and neuroscience – the emerging field of neuroeconomics – offers a potentially powerful way to study the neural mechanisms underlying reward, punishment, and decision making. In the long run, this might also advance our understanding of the dysfunction of these systems in mental disorders like schizophrenia, pathological risk-taking, and drug addiction. We contend that understanding of the neurobiological basis of decision making will be deepened by studying the economic problems solved by people and animals in their natural physical and social environments using neurophysiological, neuroimaging, and neuropharmacological techniques in the laboratory. We also suggest that studies of reward, punishment, and decision making in non-human primates will be particularly instructive, given the many biological and behavioral homologies shared by these species and humans.

In this chapter, we review the current understanding of the neural processes that construct representations of biologically-relevant value, translate these valuation signals into behavioral decisions, and update representations of value upon which future decisions are made. These results are largely derived from studies where the activity of single neurons is recorded in monkeys making choices. Our review will focus in particular on three brain regions – the orbitofrontal cortex (OFC), the lateral intraparietal area (LIP), and the posterior cingulate cortex (CGp). Value representations differ substantially in these brain areas, suggesting that distinct representation of value may contribute to distinct decision processes.

ECONOMIC CHOICE AND THE ORBITOFRONTAL CORTEX

Lesion, Neuroimaging, and Single-cell Studies of OFC and Decision Making

Consider a person sitting in a restaurant and choosing from the menu between tuna tartare and fried calamari. Presumably, the person assigns a value to the tuna and a value to the calamari, and then compares the two values to make a decision (Padoa-Schioppa *et al.*, 2006). Economic choice – the choice between different goods – thus entails assigning values to the available options. In this section, we review evidence indicating that neurons in the OFC encode the value subjects assign to the available goods during economic choice.

A strong link between economic choice and the OFC comes from observations of patients with brain damage in this area, who present choice deficits in multiple domains. For example, early clinical signs of fronto-temporal dementia (FTD, a neurodegenerative disorder that initially affects the orbital cortex and parts of the temporal lobe) include eating disorders.

In other words, FTD patients seem to assign the "wrong value" to appetitive stimuli (Pasquier and Petit, 1997). Patients with OFC lesions also display abnormal behavior in gambling tasks (Bechara *et al.*, 1996; Rahman *et al.*, 1999), suggesting a difficulty in coping with risk. However, choice deficits are also apparent in simple preference judgment tasks, in which patients with OFC lesions make inconsistent or erratic choices significantly more often than do either healthy subjects or patients with dorsolateral frontal lesions (Fellows and Farah, 2007). Finally, OFC patients also display unusual or poor choice patterns in the ultimatum game (Koenigs and Tranel, 2007) and in social contexts, as famously noted in the case of Phineas Gage (Damasio *et al.*, 1994).

A second strong link between economic choice behavior and OFC comes from imaging experiments in humans. Many studies found higher activation in the OFC when subjects were presented with pleasant stimuli in multiple sensory modalities (e.g., visual, taste, etc.) compared to neutral stimuli (O'Doherty, 2004). Most interestingly, in experiments that compared conditions in which subjects did or did not make a choice, OFC was significantly more active in the choice condition (Arana *et al.*, 2003). The same area was also more activated by high incentives compared to low incentives. In comparison, neural activation in the amygdala varied depending on the incentive level, but did not vary with task demands.

Classical results of single-cell recordings in non-human primates are also consistent with the hypothesis that OFC neurons might be a substrate for economic valuation. For example, in an early study, Thorpe and colleagues (1983) observed that neurons in OFC responded to the presentation of visual stimuli in a way that was not purely "sensory." The response of one neuron to the visual presentation of a liquid-filled syringe depended on whether in previous trials the liquid was apple juice or salted water, even though the syringe was visually indistinguishable in the two conditions. Rolls and colleagues (1989) subsequently found that the activity of OFC neurons in response to a particular taste could be altered by hunger and satiety, a modulation not observed in the primary taste area. These studies thus indicated that the activity of OFC neurons is sensitive to both the nature of physical stimulation and motivational state. More recently, Wallis and colleagues found that the activity of neurons in OFC can be modulated by the amount of juice delivered to the monkey (Wallis, 2007). Roesch and Olson (2005) observed that OFC neuronal activity varied depending on the duration of a time delay intervening before juice delivery. Interestingly, there was an inverse correlation between the effects of juice amount and the effects of time delay. Under the assumption that the neurons recorded in that study encode the subjective value at stake in any trial, one possible interpretation of this result is that the delay represents a cost for the monkey (for example due to discounting) and that OFC neurons encode net value (benefit–cost).

Neurons in OFC Encode Economic Value

Taken together, these results suggest that neurons in OFC may represent the behavioral valence of goods. However, in the experiments with monkeys described above, the animals were never asked to choose based on their own preferences. Rather, monkeys were either simply delivered juice, or they were asked to select between two options, one of which was always objectively advantageous (i.e., dominant). Consequently, these tasks could not provide a measure of the subjective value monkeys assigned to the juice. In contrast, the behavioral paradigm illustrated in Figure 29.1 provides such a measure (Padoa-Schioppa and Assad, 2006). In this experiment, thirsty monkeys chose between two different juices offered in variable amounts. When the two juices were offered in equal amounts, monkeys had a strong preference for one of the juices. However, if the other, less preferred juice was offered in sufficiently large amounts, monkeys chose it. The relative value of the two juices could thus be inferred from the indifference point – the quantity ratio at which monkeys chose either juice equally often.

In a recent study, Padoa-Schioppa and Assad used this behavioral paradigm to examine the activity of neurons in OFC as a function of subjective value (Padoa-Schioppa and Assad, 2006). Figure 29.2 illustrates the activity of one representative cell. In this session, the monkey chose between grape juice (A) and diluted cranberry juice (B). The behavioral choice pattern (Figure 29.2a, black symbols) indicated that $V(A) = 3.0\,V(B)$. The response of the neuron (Figure 29.2a, red symbols) showed a characteristic U-shape, similar to what would be expected if the neuron encoded the value chosen by the monkey in any given trial. Indeed, the activity of the neuron was low when the monkey chose 1A or when it chose 3B (in units of V(B), *chosen value* = 3); the neuronal activity was higher when the monkey chose 2A and when it chose 6B (*chosen value* = 6); the neuronal activity was highest when the monkey chose 3A and when it chose 10B (*chosen value* ≈ 10). This point can best be appreciated in Figure 29.2b, where the activity of the neuron (y-axis) is plotted directly against the variable *chosen value* (x-axis).

A further analysis confirms that U-shaped responses like these encode the subjective value monkeys assign

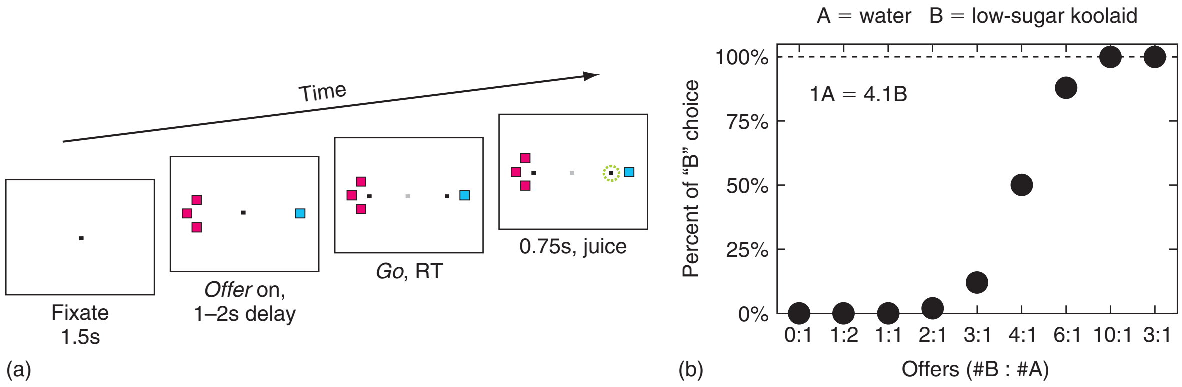


FIGURE 29.1 An economic choice task. (a) At the beginning of each trial, the monkey fixated the center of the monitor. Two sets of squares appeared on opposite sides of the fixation point (offer). Different colors of the squares indicated different juice types and the number of squares indicated the juice amount. After a randomly variable delay (1–2 s), two saccade targets appeared near the offers ("Go"). The monkey indicated its choice and maintained fixation on the saccade target for 0.75 s before juice delivery (juice). The trial was aborted if the monkey broke fixation before the Go. For any juice pair, the quantities of the two juices varied randomly. Trials with the three juice pairs were randomly interleaved and, for any given pair of offers (offer type), left/right positions were counterbalanced. (b) Choice pattern. In this session, the monkey chose between water (juice A) and low-sugar Kool-Aid (juice B). The plot shows the percentage of trials in which the monkeys chose juice B (y-axis) for various offer types (x-axis). A sigmoid fit provides the measure of the relative value 1A = 4.1B. Adapted from Padoa-Schioppa and Assad (2006, 2008).

to the juice, as opposed to any physical property of the juices. In Figure 29.2c, the activity of the same neuron is plotted as a linear function of the number of units of A (#A) and number of B (#B) chosen. The hypothesis that neuronal activity encodes the *chosen value* leads to a simple prediction regarding slopes a_A and a_B. Specifically, a_A should be proportional to the value of A, a_B should be proportional to the value of B, and the ratio $k^* = a_A/a_B$ should be equal to the value ratio V(A)/V(B). In other words, the slope ratio (k^*) provides a neuronal measure of the relative value of the two juices, independent of the behavioral measure of relative value (n^*), which represents the indifference point obtained from the sigmoid fit. If U-shaped responses indeed encode the *chosen value*, the identity $k^* = n^*$ should hold true. This condition is satisfied for the cell in Figure 2, for which the neuronal measure of relative value $k^* = 2.8(\pm 0.7)$ is statistically indistinguishable from the behavioral measure $n^* = 3.0$. In fact, the identity $k^* = n^*$ holds true in general. Considering the entire population, a linear regression provides $k^* = -0.13(\pm 0.15) + 1.05(\pm 0.15)n^*$ (average across 10 juice pairs), which is statistically indistinguishable from $k^* = n^*$. This result demonstrates that U-shaped responses indeed encode value as a subjective quantity, as opposed to any physical property of the juices. Indeed, if U-shaped responses encoded a physical property of the juices (e.g., sugar content), U-shapes should not vary from session to session depending on the relative value the monkey assigns to the juices in that particular session. In other words, k^* should be independent of n^* across behavioral sessions, contrary to what is observed. Hence, U-shaped responses encode the value monkeys assign to the juice they choose to consume (Roesch and Olson, 2005).

Neurons encoding the *chosen value* were frequent in OFC, but other types of neuronal responses were also found. For example, OFC neurons often encoded the *offer value* – that is, the value of one of the two juices alone. Figure 29.3 (a and b) shows the activity of two neurons encoding, respectively, *offer value A* and *offer value B*. Other frequently observed responses varied in a binary fashion depending on the type of juice chosen by the monkey, independently of the amount (Figure 29.3c). These responses appear to encode the juice *taste*.

One fundamental aspect of the value representation in OFC is that neuronal responses in this area do not depend on the visuo-motor contingencies of choice. This point can be observed in Figure 29.2d, where the activity of the same neuron shown in Figure 29.2a is plotted separately for trials in which the monkey indicated its choice with an eye movement to the left (red) or to the right (green). The activity was nearly identical for the two groups of trials. In other words, this neuron encoded the value of the chosen juice *per se*, independently of the action used by the monkey to indicate its choice. Similarly, the activity of this cell did not depend on the spatial configuration of the stimuli on the monitor (i.e., whether juice A was offered on the right and juice B on the left, or *vice versa*). The same was true across the population. For more than 95% of neurons in OFC, the activity was independent of the visuo-motor contingencies of choice.

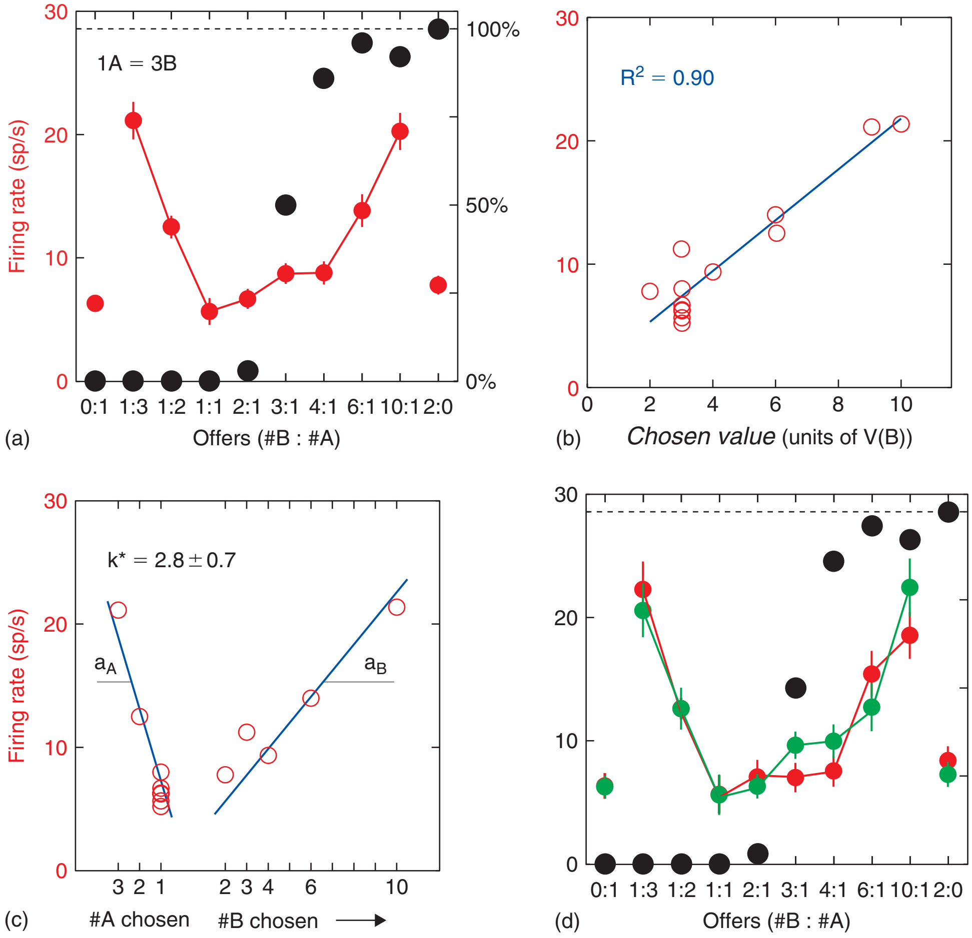


FIGURE 29.2 Example OFC neuron encoding the *chosen value*. (a) Black symbols represent the behavioral choice pattern and red symbols represent the neuronal activity (±s.e.m.). We conventionally express values in units of V(B). The cell activity is low when the monkey chooses 1A and when it chose 3B (offer types 1B:1A, 2B:1A, 3B:1A; *chosen value* = 3); it is higher when the monkey chooses 2A and when it chose 6B (offer types 1B:2A, 6B:1A; *chosen value* = 6); and it is highest when the monkey chooses 3A and when it chose 10B (offer types 1B:3A, 10B:1A; *chosen value* ≈ 10). (b) Same neuronal responses plotted against the variable *chosen value* (expressed in units of V(B)). A linear regression provides R2 = 0.90. (c) Same neuronal response plotted against the number of A and B chosen (linear scale). The slope ratio *k** provides a neuronal measure of relative value statistically indistinguishable from the behavioral measure. (d) Here, the cell activity is plotted separately for trials in which the monkey indicated its choice with an eye movement to the left (red) or to the right (green). The neuronal responses recorded for the two sets of trials are essentially identical, indicating that the activity of this neuron was independent of the visuomotor contingencies of choice. This was consistently true for over 95% of neurons in OFC. Adapted from Padoa-Schioppa and Assad (2008).

As described in the following sections, the observation that neurons in OFC represent value independently of the sensory-motor contingencies of choice fundamentally distinguishes this representation of value from that found in other brain areas in monkeys. Interestingly, this characteristic also seems to distinguish OFC responses in primates and rodents. Indeed, two recent studies found that neuronal activity in the rat OFC actually does depend on sensory and motor contingencies (Feierstein *et al.*, 2006). One possibility is that the region examined in rats was not homologous to the region examined in monkeys. Alternatively, it is possible that an abstract representation of value removed from sensory and motor contingencies emerged subsequent to the divergence of primates and rodents from a common ancestor (Roesch *et al.*, 2006).

The Representation of Value in OFC is Invariant for Changes of Menu

For any neuronal representation of value, a broad and fundamental question is whether and how value-encoding neuronal responses depend on the behavioral context. In principle, two ways in which the

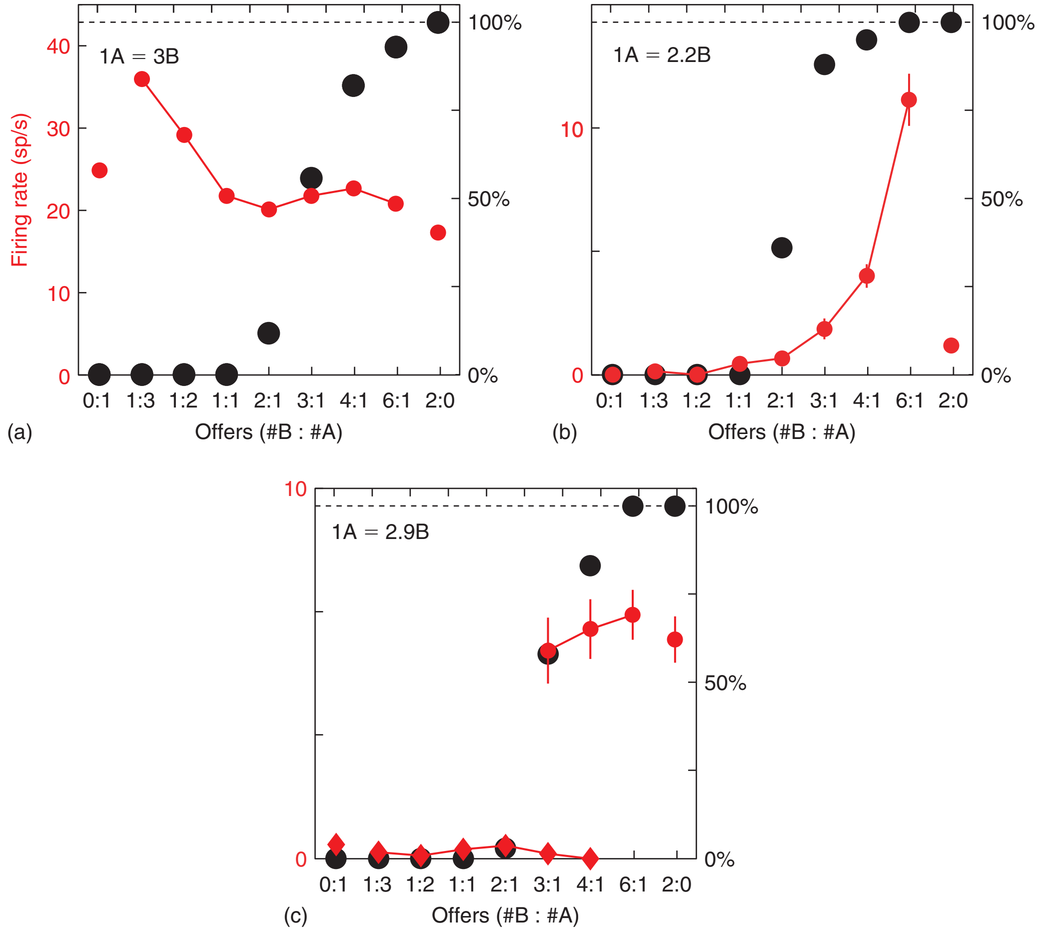


FIGURE 29.3 Different types of neuronal response patterns in OFC. Activity of three neurons encoding the variables (a) *offer value A*, (b) *offer value B*, and (c) taste. All conventions are as in Figure 29.2. Adapted from Padoa-Schioppa and Assad (2006) and Padoa-Schioppa (2007).

behavioral context might change can be distinguished. Changes of "menu" are moment-to-moment changes in the available options. For example, if a person in a wine store compares serially different pairs of bottles, changes from one pair-wise comparison to the next are changes of menu. From a computational point of view, a valuation system invariant for changes of menu presents clear advantages, because it guarantees stable choices and preference transitivity (see Box 29.1). In contrast, changes of "condition" (including changes of the chooser's internal state) are changes that occur on a longer timescale. For example, if that same person goes from the wine store (where different options cost about $15) to a car retailer (where different options cost $15,000 or more), his valuation system might adapt to the new range of values. Similarly, the valuation system could also adapt to changes of internal motivation.

In a second study, Padoa-Schioppa and Assad (2008) examined specifically whether the representation of value in OFC depends on the menu – that is,

BOX 29.1

TRANSITIVITY

Transitivity is a fundamental trait of economic choice behavior. Given three options, X, Y, and Z, if an individual prefers X to Y and Y to Z, she ought to prefer X to Z (*preference transitivity*). Likewise, if an individual is indifferent between X and Y and between Y and Z, she ought to be indifferent between X and Z (*indifference transitivity*). We indicate with $n_{X:Y}$ the relative value of X and Y, such that $V(X) = n_{X:Y}\ V(Y)$. Under the assumption of linear indifference curves (i.e., if value functions for different goods are all the same up to a scaling factor), indifference transitivity reduces to the relationship $n_{A:B} = n_{B:C} * n_{A:C}$ (*value transitivity*) (Padoa-Schioppa and Assad, 2008).

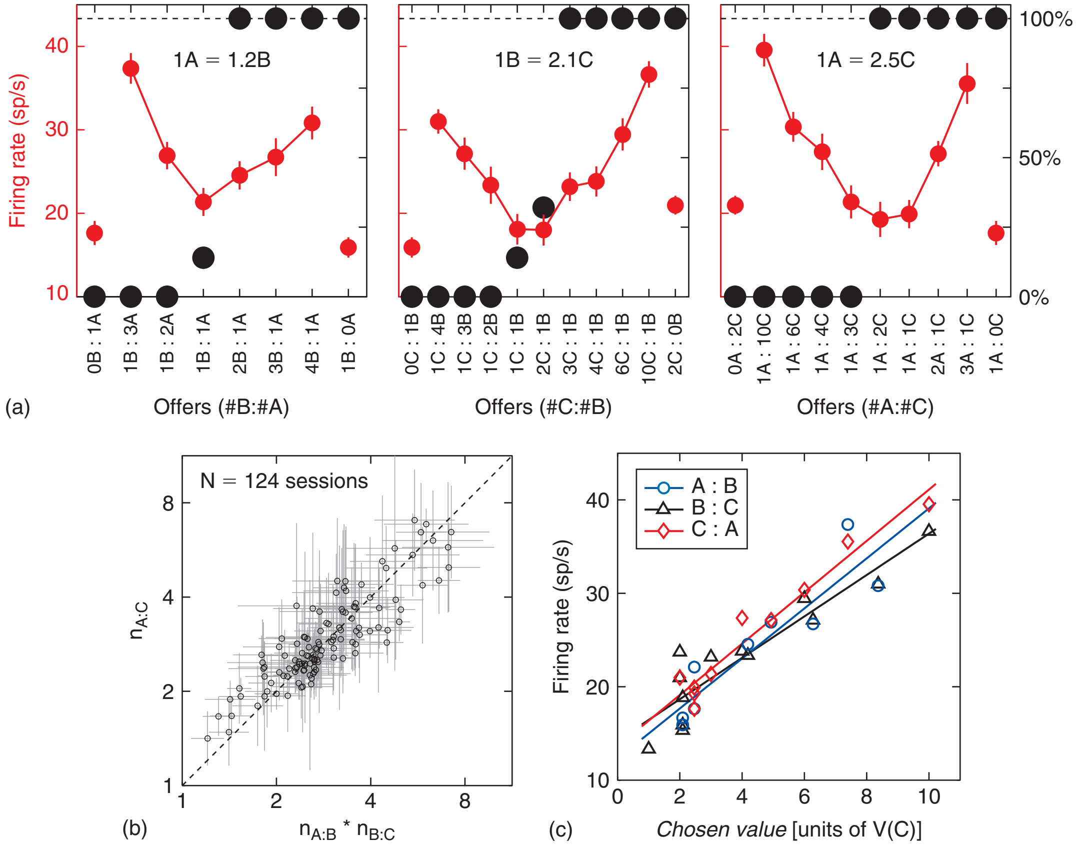


FIGURE 29.4 Menu invariance in OFC. (a) Responses of one neuron encoding the *chosen value*. The three panels refer to trials A:B, B:C, and C:A. In each panel, black symbols represent the behavioral choice pattern and red symbols represent the neuronal activity (±s.e.m.). Relative values (top left) combine according to transitivity (because $1.2 * 2.1 \approx 2.5$). The neuronal response encodes the *chosen value* independently of the juice pair. (b) Value transitivity across sessions. Each data point in the scatter plot represents one session, and all 124 sessions are shown. For each session, gray error bars represent the errors of measure (s.d.). The diagonal dashed line corresponds to $n_{A:C} = n_{A:B} * n_{B:C}$. Relative values measured in any given session satisfy transitivity unless they are significantly removed from this line. Value transitivity was satisfied by all 124 sessions ($P > 0.05$). (c) Same neuronal response as in (a), combining data from the three juice pairs. The firing rate (y-axis) is plotted against variable chosen value, and different symbols and colors refer to the three juice pairs (blue circles for A:B, black triangles for B:C, and red diamonds for C:A). Each symbol represents one trial type. The regression lines are obtained from a full-model analysis of covariance (ANCOVA). Adapted from Padoa-Schioppa and Assad, (2008).

whether neuronal responses encoding the value of one particular good depend on what other goods are available at the same time. In this case, monkeys chose between three juices (A, B, and C, in decreasing order of preference). In each trial, monkeys chose between two juices, and trials with the three juice pairs (A:B, B:C, and C:A) were randomly interleaved. For example, in the session illustrated in Figure 29.4a, a monkey chose between grape juice (A), fruit punch (B), and apple juice (C). The three panels in the figure refer, respectively, to trials A:B, B:C and C:A, and black symbols represent the choice patterns. Notably, the three relative values obtained from the sigmoid fits combine according to value transitivity, in the sense that $1.4 * 1.9 \approx 2.6$. The values monkeys assigned to different goods satisfied transitivity in general (Figure 29.4b). As a consequence, quantities of the three juices could be expressed on the same value scale.

Padoa-Schioppa and Assad found that neuronal responses in OFC were typically invariant for changes of menu. For example, the activity of the cell illustrated in Figure 29.4a (red symbols) encoded the *chosen value* independently of the juice pair. This can be observed most clearly in Figure 29.4c, where the same neuronal response (y-axis) is plotted against the variable *chosen value* (x-axis) and the three colors represent the three juice pairs. The three regression lines (from an analysis of covariance) are statistically indistinguishable from one another, indicating that the activity of the cell encodes the *chosen value* independently of the menu. Similarly, the activity of neurons encoding the *offer value* or the taste of one particular juice did not typically depend on the other type of juice offered in the same trial. Across a population of 868 neuronal responses[1], 767 (88%) were invariant for changes of menu.

Whether the activity of neurons in OFC is causally related to economic choice remains an open question. Nonetheless, it is worth noting that neurons in OFC encoding economic value independently of other options available could in principle provide a neuronal explanation for preference transitivity. Indeed, in most typical situations, human and animal economic choices satisfy preference transitivity (Mazur and Coe, 1987). When transitivity violations are observed (Tversky, 1969; Navarick and Fantino, 1972; Shafir, 1994), it is generally because assigned values depend on the menu (Grace, 1993; Shafir, 1994). In other words, menu-invariant values imply preference transitivity. If assigned values (as measured at the behavioral level) result from the activity of neurons in the OFC, then menu-invariant responses in OFC imply menu-invariant values, which in turn imply transitive preferences. As described below, menu invariance may be a trait specific to OFC. Indeed, in other brain areas where this issue has been examined, signals reflecting the value of one particular option are modulated by the value of other options available at the same time.

Ordinality, Cardinality and Neuronal Adaptation

At first glance, the results illustrated in Figure 29.4 seem to differ from the results of a previous study by Tremblay and Schultz (1999). In their experiment, these authors delivered to monkeys one of three types of juice (A, B, and C, in decreasing order of preference) in fixed amounts. Trials were blocked, with one pair of juices employed in each block. Tremblay and Schultz found OFC neurons that responded to juice A but not to juice B during "A:B" blocks, and to juice B but not to juice C in "B:C" blocks. They interpreted these results to suggest that OFC neurons encode the "relative preference" of the juices.

Notably, the apparent discrepancy between the experimental results obtained in the two studies (Padoa-Schioppa and Assad, 2008; Tremblay and Schultz, 1999) leads to very different hypotheses regarding the nature of the encoding in OFC. On the one hand, the results of Tremblay and Schultz suggest that the OFC neurons encode the relative preference (i.e., the ordinal ranking) of the juices. On the other hand, the results of Padoa-Schioppa and Assad suggest that OFC neurons encode value in a cardinal (i.e., number-like) sense. From the point of view of economic theory, the issue of ordinal versus cardinal utility is foundational (Kreps, 1990).

How can the two sets of results be reconciled? One possibility is that recordings in the two studies examined different brain regions. Although the anatomical reconstructions in the two studies do not present clear differences, a precise distinction between different orbital areas requires refined histological procedures (Carmichael and Price, 1994) that were not conducted in either study. The difference between the behavioral paradigms used in the two studies suggests an alternative hypothesis. Specifically, it is possible that the observations of Tremblay and Schultz critically depended on the fact that trials were presented in blocks. In principle, a block design could affect neuronal responses in multiple ways. For example, in their study, "A:B" blocks could be considered high-value blocks, whereas "B:C" blocks could be considered low-value blocks. Their observations thus suggest that the activity of OFC neurons might adapt to the general behavioral context defined across many trials (the behavioral "condition", as defined above). In this view, the changes in neuronal activity observed by Tremblay and Schultz were not due to menu dependence, but rather to a slowly adapting neuronal representation.

Taken together, the two sets of results thus suggest that neurons in OFC encode value in a cardinal sense (as opposed to ordinal preference), and that this neuronal representation of value adapts to the behavioral condition. Critically, whether the neuronal population examined by Padoa-Schioppa and Assad indeed undergoes such adaptation remains to be established.

VALUATION SIGNALS IN PARIETAL CORTEX

Value Modulates Decision-related Signals in Parietal Cortex

In contrast with the OFC, parietal cortex appears to play an important role in linking sensory signals with motor commands, as well as guiding sensory attention, based on evidence from neurological, neurophysiological, and neuroimaging studies (Gnadt and Anderson, 1988; Colby *et al.*, 1996). The importance of action for biological fitness implies that sensory-motor processing and attention should incorporate the value of alternative interpretations of sensory data for guiding behavior. By scaling neuronal activity that links

---

[1] A "neuronal response" is defined as the activity of one neuron in one time window.

sensation to action by value, motor systems may be biased to generate actions with greater behavioral utility. Similarly, scaling attention by the potential value of different stimuli may enhance the ability to detect and discriminate objects and events of high behavioral utility.

These ideas were first tested in a series of studies by Platt and Glimcher (1999) in which they explicitly probed the modulation of neuronal activity in the lateral intra-parietal area (LIP) by the expected value of available options. Expected value *E(x)* is defined as the product of expected gain, *x*, and the likelihood that gain will be realized (Arnaud and Nichole, 1982 (1662)). Prior studies had demonstrated that LIP neurons respond to visual stimulation as well as preceding gaze shifts to visible and remembered target locations (Gnadt and Andersen, 1988; Goldberg *et al*., 1990). Moreover, LIP neurons appear to signal the relative importance of visual stimuli for guiding subsequent behavior (Colby *et al*., 1996; Platt and Glimcher, 1997). Such observations suggested the hypothesis that LIP links sensation to action according to the expected value of each possible response.

Platt and Glimcher first tested this hypothesis in a task in which monkeys were not permitted to make choices, thus permitting the authors to vary the expected value of each option independently of potentially confounding factors associated with active decision making. In this task, monkeys were cued by the color of a fixation stimulus to shift gaze to one of two peripheral visual targets, while the activity of single LIP neurons was monitored. The expected value of shifting gaze to each target was systematically varied by either delivering different amounts of fruit juice for correct gaze shifts to each of the targets, or by altering the probability that each of the possible gaze shifts would be cued across blocks of trials. In both cases, the authors found that when cue color, target location, and movement metrics were held constant, the activity of many neurons in area LIP was proportional to the expected value of a specific target. Similar correlations between neuronal activity and the expected value of a particular movement (either movement probability or expected reward magnitude) have been found in prefrontal cortex , the caudate nucleus and substantia nigra pars reticulate of the basal ganglia, and the superior colliculus (Salzman *et al*., 2005; Hikosaka *et al*., 2006). In a second study, Platt and Glimcher further demonstrated that, in the absence of an overt cue indicating which movement would be rewarded, the frequency with which monkeys chose each target was proportional to its expected value. Moreover, the activity of many LIP neurons paralleled these value-based decisions and was a direct function of target preference.

Taken together, these studies indicate that brain areas implicated in the conversion of sensory stimuli into action, such as LIP, incorporate the value of each available option. In Platt and Glimcher's studies, however, the value of the available options remained constant throughout blocks of 50–100 trials, thus making it difficult to discern how closely LIP neurons track local fluctuations in value associated with the dynamics of ongoing decision making. Sugrue, Corrado, and Newsome (Sugrue *et al*., 2004) extended these observations by probing decision-related activity in LIP using a foraging task (Newsome *et al*., 2008). In their study, the likelihood of rewards associated with each of two targets fluctuated over time depending on the monkeys' recent choices. Under these conditions, monkeys tended to match the rate of choosing each target to the relative rate of reinforcement of that target over both short and long timescales (matching behavior). Moreover, the responses of individual LIP neurons to a particular target corresponded to the history of relative payoffs associated with each target, with the greatest weight placed on the most recent trials. Similar results were found by Dorris and Glimcher (2004) in monkeys performing a frequency-dependent foraging task. In their study, LIP neurons were found to reflect a "value weight": the activity of each neuron was modulated by the value of the corresponding visual stimulus divided by the value sum of all visual stimuli. Together, these and other studies suggest that behavioral decisions may be computed by scaling neuronal responses to sensory stimuli and motor plans by their expected value, thus modulating the likelihood of reaching the threshold for generating a particular percept or eliciting a specific action (Gold and Shadlen, 2001).

Valuation Signals in Parietal Cortex are Independent of Modality

Although concrete outcomes such as eating, drinking, or sex clearly motivate behavior, abstract goals such as information gathering or social interaction can also motivate behavior in the absence of hedonic experience and thus should contribute to the value of any potential action. For group-living species such as humans and many non-human primates, the social environment strongly influences the behavioral context in which individuals pursue rewards, avoid punishments, evaluate risks, and make decisions. The adaptive significance of navigating a complex social environment suggests that social stimuli might evoke neural activity in some of the same circuits that process primary rewards and punishments, and subsequently modulate the neural valuation functions that guide attention and action. For example, male primates use visual cues to predict female mating receptivity (Hrdy and Whitten, 1987) and field studies show that monkeys preferentially invest in relationships with dominant individuals (Cheney and Seyfarth, 1990). These observations suggest that the primate brain also computes valuation functions for specific social and reproductive stimuli that guide adaptive behavior.

These observations led Platt and colleagues to hypothesize a neural system linking social stimuli, such as images of faces or bodies, to the valuation functions guiding action. Deaner and colleagues (2005) explored this hypothesis behaviorally using a "pay-per-view" task in which thirsty male rhesus macaques were given a choice between two visual targets. Orienting to one target yielded fruit juice; orienting to the other target yielded fruit juice and the picture of a familiar monkey. By systematically changing the juice amounts for each target as well as the picture shown, the authors estimated the value of different types of social and reproductive stimuli in a liquid currency. Their study revealed that male monkeys forego larger juice rewards in order to view female sexual signals or faces of high-ranking males, but require these large rewards to view the faces of low-ranking males and females (Figure 29.5a). Hayden and colleagues (2007) extended these findings by demonstrating that humans, like monkeys, will also pay more to view pictures of attractive members of the opposite sex than to view pictures of unattractive ones, even when the reward cues are implicit. Specifically, men placed a value of around half a cent (US) on the opportunity to view an attractive woman, whereas the value women placed on the opportunity to view an attractive man was not different from zero (Figure 29.5b).

These findings suggest that decisions based on value operate on a common currency that is independent of the modality of the goods under consideration or the actions they motivate. When monkeys choose between fluid and social rewards, they show consistent, apparently adaptive, preferences. Likewise, human subjects systematically trade off monetary and pictorial rewards. These observations provoke the hypothesis that the brain transforms information about disparate options into a common currency of value in which these options can be compared and evaluated. The studies described above suggest that the OFC encodes the abstract value of goods under consideration. These behavioral and neurobiological observations predict, then, that modulation of sensory-motor processing in cortical areas like LIP, which presumably lie downstream of abstract value processing in OFC, should be independent of the modality of the desired outcome. That is, it should not matter to an LIP neuron whether the option in its receptive field is rewarding because it is associated with juice, money, or the opportunity to look at an attractive member of the opposite sex – as long as revealed preferences indicate these goods have the same subjective value. Since the goal of action is presumably to maximize behavioral utility, sensory-motor decision processes should be modulated by value independent of the modality of the outcome.

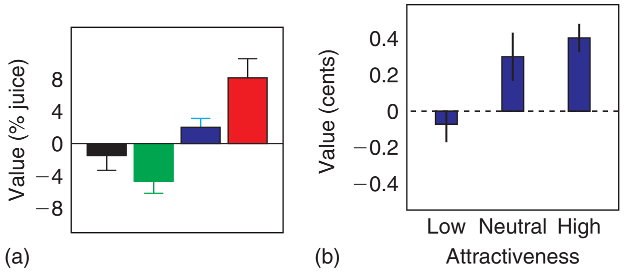


FIGURE 29.5 Monkeys and people value visual social information. (a) Average value (in percent juice volume) of the opportunity to view an image of a gray square (gray), subordinate male face (green), dominant male face (blue), or female perineum (red), for five male macaques. Value was estimated using a "pay-per-view" task in which monkeys chose between fluid rewards and fluid rewards paired with images. (b) Average value (in cents) of viewing the face of a female of low, neutral, or high physical attractiveness, for 20 male subjects. Value was estimated using a "pay-per-view" task in which people chose between monetary rewards and monetary rewards paired with images.

Platt and colleagues tested this idea directly by examining the activity of LIP neurons in monkeys performing the pay-per-view task described above (Klein *et al.*, 2008). In this study, the target associated with the display of an image was positioned within the receptive field of a neuron under study, whereas the other target was positioned in the other visual hemifield. Across blocks of trials, the identity of the class of images displayed for choosing the target in the receptive field was varied, and the volume of juice delivered for choosing either target was also varied. The authors found that LIP neurons were sensitive to both visual reward outcomes and juice reward outcomes associated with choosing the target in the neuronal receptive field (Figure 29.6a). Specifically, modulation of neuronal activity matched the value monkeys placed on seeing particular classes of images, in addition to the size of juice rewards; firing rates were highest when monkeys chose to view images of female reproductive areas, slightly lower when monkeys chose to view the faces of dominant males, and lowest when monkeys chose to view the faces of subordinate monkeys. Most importantly, LIP

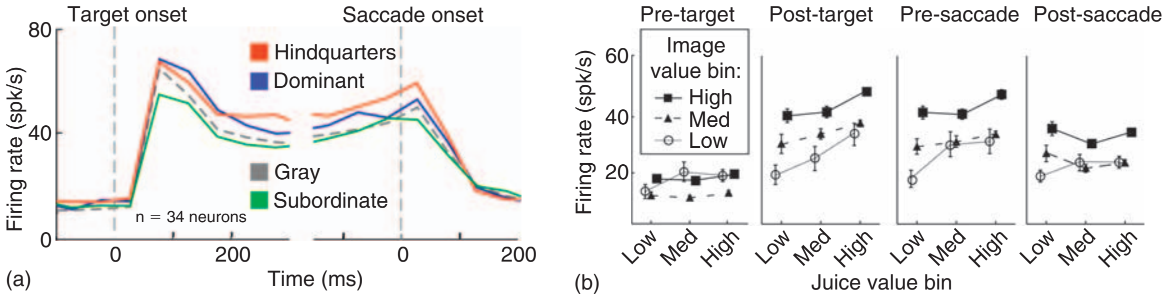


FIGURE 29.6 LIP neurons encode the abstract value of orienting to a visual stimulus. (a) Peri-stimulus time histogram (PSTH) plotting firing rate of a population of neurons as a function of time, aligned on either target onset or the saccade monkeys used to report their preference. On each trial, the monkeys chose to view an image associated with a target in the response field of the neuron under study. Images were female perinea (red), dominant male faces (blue), subordinate male faces (green), or gray squares (gray). (b) LIP neurons encode both fluid value and juice value. Firing rate is plotted in four consecutive 200-ms bins, and sorted by juice value and image value. ANOVA for juice value and image value were significant, but there was no interaction between these variables.

neurons encoded the contributions of expected visual outcomes and expected fluid outcomes to target value independently (Figure 29.6b). Thus, LIP neurons appear to signal the value of a visual target derived from the multiple potential outcomes, either visual or fluid, that could occur. Importantly, value modulation was not observed when monkeys were not permitted to choose where to look and were forced to make a particular behavioral response. This result is consistent with the idea that LIP neurons signal the relative value of the options available for orienting (Dorris and Glimcher, 2004; Sugrue *et al.*, 2004) or the likelihood that they will look towards a particular target (Gold and Shadlen, 2001).

These observations indicate that value modulation of sensory-motor processing in parietal cortex, and presumably other areas that contribute to decision making, is relative and not absolute, in contrast with value signals in OFC. Consistent with this notion, Platt and colleagues found that LIP neurons are also sensitive to the number of targets presented within their receptive fields (Roitman *et al.*, 2007). In that study, monkeys were simply rewarded for orienting to a single target opposite the receptive field of the neuron under study. While monkeys waited to make the required movement, an array of dots was illuminated within the receptive field. The firing rates of about half of the neurons studied were systematically suppressed as the number of elements in the array increased (Figure 29.7). Similar findings have been reported for neurons in the superior colliculus by Basso and Wurtz (1997).

These data suggest that LIP neuronal responses to stimuli and their associated actions may be normalized by the number of options available, similar to the process of divisive normalization characteristic of neurons in primary visual cortex (Heeger, 1993; Schwartz and Simoncelli, 2001). Since expected value depends on both the magnitude of expected reward and the likelihood that the reward will be delivered, normalization by the number of possible targets effectively rescales target valuation signals on a relative scale. Thus, neurons in parietal cortex, and most likely other areas that convert sensory information into action, appear to signal the relative value of choosing one of the available options for control of behavior.

Valuation, Attention, and Decision Making

Most studies of the representation of value by neurons in the primate brain have required monkeys to express their choices by orienting the eyes to a visual stimulus. Such studies have revealed that increasing the value of orienting to visual targets is associated with enhanced neuronal activity in a number of areas implicated in visual orienting behavior, including LIP (Platt and Glimcher, 1999; Coe *et al.*, 2002; Sugrue *et al.*, 2004), dorsolateral prefrontal cortex (Leon and Shadlen, 1999), the supplementary eye-fields (Amador *et al.*, 2000), the caudate nucleus (Kawagoe *et al.*, 1998; Watanabe *et al.*, 2003), substantia nigra pars reticulate (Sato and Hikosaka, 2002), anterior (Ito *et al.*, 2003) and posterior (McCoy *et al.*, 2003) cingulate cortices, and the superior colliculus (Ikeda and Hikosaka, 2003). Value modulation in these areas has been interpreted to reflect decision-related computations that bias orienting to a particular target (Platt and Glimcher, 1999; Gold and Shadlen, 2001; Sugrue *et al.*, 2004). However, previous studies have also shown that neurons in most of these areas are sensitive to attention (Goldberg *et al.*, 1990; Colby *et al.*, 1996; Gottlieb *et al.*, 1998), and lesions to many of these areas are often associated with deficits in attention. Since attention is typically controlled in animal studies by manipulating rewards, and rewarding events might also draw

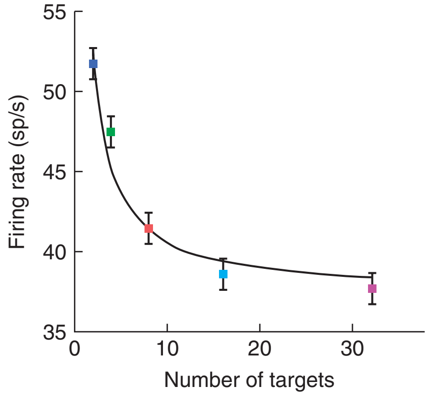


FIGURE 29.7 Normalization of neuronal activity in LIP by target number.An array of irrelevant dots was displayed in the response fields of neurons in LIP while monkeys waited to shift gaze in the opposite direction. For a subpopulation of neurons, firing rate systematically decreased with increasing number of dots. Divisive normalization by target number may contribute to relative value scaling of target-related activity in LIP.

attention (Horvitz, 2000), it is possible that reward modulation of neuronal activity in brain areas associated with visual orienting might actually reflect changes in attention, and *vice versa* (Maunsell, 2004). In other words, attention and value are naturally linked and thus difficult to disentangle either behaviorally or neurophysiologically.

In an initial attempt to address this issue, Bendiksby and Platt found that activation of LIP neurons was enhanced simply by increasing the value of all potential targets and that this enhancement was associated with predictable changes in behavioral response time (Bendiksby and Platt, 2006). In their study, monkeys performed a peripheral attention task in which they were rewarded for shifting gaze to a response target when one of two eccentric cues briefly flickered. The cues were presented sequentially, and flicker duration was titrated so that monkeys would perform correctly on about 70% of trials. The first cue to be illuminated was 80% likely to flicker, while the second was only 20% likely to flicker. Thus, monkeys should have paid closer attention to the first cue illuminated. Importantly, monkeys were uniformly rewarded with large or small squirts of juice in separate blocks of trials for correct performance – independent of which cue flickered. Under these conditions, monkeys responded more quickly when they could expect larger rewards than when they could expect smaller rewards, suggesting that action-encoding neural signals reached the threshold for movement initiation more rapidly when monkeys were more motivated (Figure 29.8a). At the same time, LIP neurons responded more strongly to visual cues in their receptive fields in large reward blocks than in small reward blocks, and these value-related modulations persisted until the time of the behavioral response (Figure 29.8b).

These results indicate that LIP neurons are sensitive to the value associated with visual stimuli, even when the stimulus is divorced from the ensuing behavioral response. Moreover, increasing stimulus value appears to increase both the gain of the initial neuronal response, as well as the sustained level of neuronal activity following stimulus presentation. Balan and Gottlieb have reported analogous gain modulation and baseline activity changes in LIP when task difficulty and attention, respectively, were manipulated (Balan and Gottlieb, 2006). Together, these observations suggest that attention and value jointly determine sensory-motor processing in LIP, and possibly other areas that connect sensation and action. More research is needed to determine precisely how value and attention influence information processing by neurons in different brain areas during decision making.

EVALUATION OF EVENTS AND ACTIONS IN POSTERIOR CINGULATE CORTEX

Anatomical, Clinical, and Neuroimaging Significance of Posterior Cingulate Cortex

Evidence reviewed in this chapter suggests that the abstract value of goods under consideration for choice is represented in the OFC. Translation of abstract valuation signals represented in OFC, and possibly elsewhere, into action appears to involve value-related modulation of cortical and sub-cortical areas involved in sensory-motor integration. One important question is how the abstract value signals in OFC, and other areas, are evaluated and bound to events and actions in order to adaptively influence future behavior.

Based on the anatomy, one potential candidate is the posterior cingulate cortex (CGp). Posterior cingulate cortex is strongly interconnected with brain areas known to be involved in learning and motivation or that are sensitive to reinforcement contingencies, including the anterior and lateral thalamic nuclei (Gabriel *et al.*, 1991), the caudate nucleus (Powell, 1978; Yeterian and Van Hoesen, 1978; Baleydier and Mauguiere, 1980), and medial portions of the OFC (Baleydier and Mauguiere, 1980). In addition, CGp is strongly interconnected with anterior cingulate cortex, which contains neurons carrying nociceptive (Sikes and

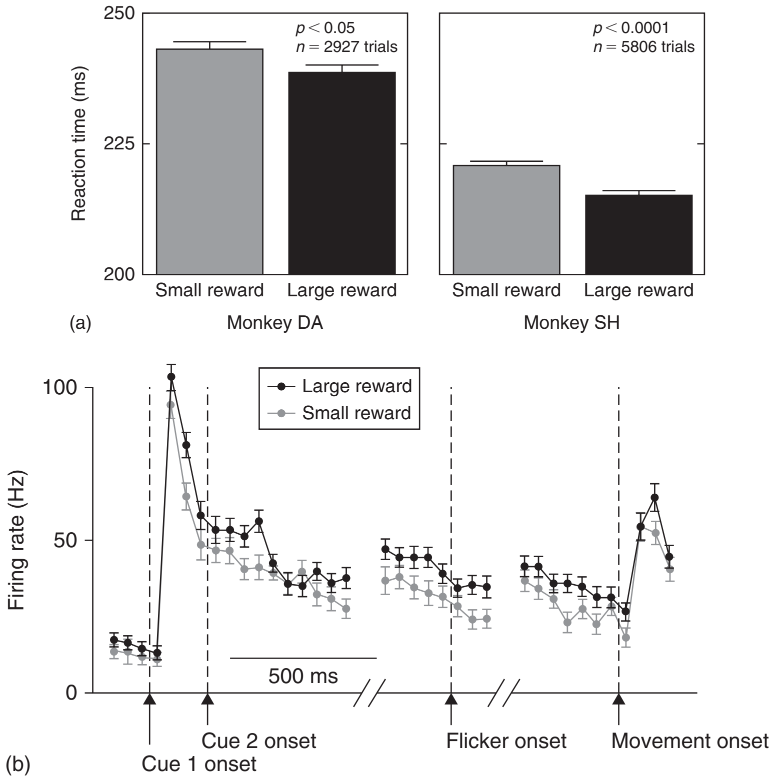


FIGURE 29.8 Increasing target value reduces reaction time and increases the gain of visual responses in LIP. Monkeys performed a peripheral attention task and earned large and small juice rewards for correct performance in different blocks of trials. (a) Reaction times for two monkeys as a function of reward size. (b) PSTH for a single LIP neuron plotting firing rate as a function of time, aligned on onset of the visual cue in the response field (left), discriminative event (middle), or gaze shift away from the response field (right). Both the gain of the visual response, and sustained activity, increased with increasing target value, independent of the movement away from the response field.

Vogt, 1992) and reward-related information (Niki and Watanabe, 1979; Shidara and Richmond, 2002; Ito *et al.*, 2003) and which is capable of activating brain reinforcement circuits when artificially activated (Goodall and Carey, 1975; Spence *et al.*, 1985). CGp is also particularly well-situated anatomically to receive information about visual events, action, and attention, including strong connections with parietal cortex (Baleydier and Mauguiere, 1980; Pandya *et al.*, 1981; Vogt and Pandya, 1987; Cavada and Golman-Rakic, 1989a, 1989b; Andersen *et al.*, 1990; Blatt *et al.*, 1990; Morecraft *et al.*, 1993), dorsolateral prefrontal cortex (Barbas and Mesulam, 1985; Selemon and Goldman-Rakic, 1988; Barbas and Pandya, 1989), and the frontal eye-fields (Barbas and Mesulam, 1981; Vogt and Pandya, 1987). CGp thus receives appropriate inputs to signal the motivational significance of visual events and action.

Consistent with its anatomical connections, clinical evidence suggests that posterior cingulate cortex contributes to visual orienting and navigation, attention, emotion, and learning and memory (Maddock, 1999). The first manifestations of Alzheimer's disease, learning and memory impairment, have been linked to a less than 20% decrease in the metabolic activity of posterior cingulate cortex (Minoshima *et al.*, 1997). Damage to CGp is also associated with spatial disorientation and navigational impairments (Cammalleri *et al.*, 1996; Takahashi *et al.*, 1997; Katayama *et al.*, 1999). Posterior cortical atrophy (or PCA), which is association with depression of CGp metabolic activity, is characterized by optic ataxia and disturbances of visual attention (Hof *et al.*, 1993, 1997; Fletcher, 1994; Braak *et al.*, 1996). Conversely, elevated activity in posterior cingulate cortex is found in a variety of personality, mood, and anxiety disorders, including schizophrenia, major depression, obsessive-compulsive disorder, and social phobia (Maddock, 1999), all of which can be viewed to some degree as disorders of behavioral valuation.

Additionally, neuroimaging studies demonstrate that posterior cingulate cortex is activated by changes in subjective motivational state, as well as by attention to events and actions. For example, Small and colleagues (2001) tested human subjects before and after feeding them with chocolate to satiation. Intriguingly, BOLD signal in CGp was elevated when subjects rated chocolate as either highly pleasant or highly unpleasant, rather than neutral. Maddock and colleagues (2003) also demonstrated activation of CGp by presentation of both positive and negative emotion-laden words, as well as retrieval of words presented in negative emotional contexts (Maratos *et al.*, 2001). Activation in CGp has also been linked to errors in reward prediction during risky decision making (Dickhaut *et al.*, 2003). More recently, Kable and Glimcher (2007) demonstrated that CGp activation varies systematically with time-discounted value in an inter-temporal choice task. Functional imaging studies have also revealed CGp activation following illumination of visual stimuli (Yamasaki *et al.*, 2002) following a shift in visual attention (Kim *et al.*, 1999; Hopfinger *et al.*, 2000, 2001; Mesulam *et al.*, 2001; Small *et al.*, 2003), during overt visual orienting (Berman *et al.*, 1999), and during visuospatial navigation (Flitman *et al.*, 1997; Ghaem *et al.*, 1997; Pine *et al.*, 2002). Moreover, the BOLD signal in CGp correlates with improvements in visual detection performance associated with valid spatial cues in peripheral attention tasks in humans (Small *et al.*, 2003). Together, these observations suggest a role for CGp in signaling motivationally significance events and actions, as well as perhaps their subjective value for guiding future behavior.

Neurophysiological Evidence that CGp Plays a Role in Evaluating Behavioral Events

Neurophysiological studies conducted in animals support the idea that posterior cingulate cortex contributes to the evaluation of stimuli as well as behavioral actions. For example, CGp neurons respond to visual stimulation (Kalia and Whitteridge, 1973; Olson and Musil, 1992; Olson *et al.*, 1996) and their responses depend on behavioral context (Dean *et al.*, 2004a). Specifically, CGp neurons respond most strongly to visual events that are unpredictable in space or time, and the strength of these neuronal responses predicts how accurately monkeys subsequently orient (Dean *et al.*, 2004). CGp neurons respond strongly after orienting movements, and these responses are anchored to locations in space rather than to the locus of retinal stimulation (Dean *et al.*, 2004a; Dean and Platt, 2006). CGp neurons also respond following the delivery of unpredictable rewards as well as following the omission of predictable rewards (McCoy *et al.*, 2003). Moreover, the CGp responses to task-related events are modulated by their associated value (McCoy *et al.*, 2003). Thus, CGp neurons appear to report potentially significant events in a manner divorced from pure sensory properties or the actions these events might guide. CGp thus carries information that could in principle be used to link events and outcomes in a context-dependent fashion. Such information might inform sensory-motor processing in brain areas, such as LIP, that contribute to action selection (McCoy *et al.*, 2003; Dean *et al.*, 2004b).

Evaluative Signals in CGp are Subjective

It is well known that choosers demonstrate idiosyncratic attraction or aversion to options associated with risky payoffs. Thus, risk sensitivity provides a powerful assay to dissociate subjective representation of value or utility from the representation of objective rewards. Recent neuroimaging studies in humans have revealed that preference for a risky option is associated with increases in neuronal activity in the ventral striatum and posterior parietal cortex (Kuhnen and Knutson, 2005; Huettel *et al.*, 2006; Preuschoff *et al.*, 2006). Moreover, choosing a risky option activates the dorsal striatum, posterior cingulate cortex, and precuneus (Dickhaut *et al.*, 2003) as well as amygdale (Hsu *et al.*, 2005) and insula (Huettel *et al.*, 2006).

These observations suggest that the neural mechanisms that bind value to events and action must also be sensitive to risk if these processes contribute to choice. To test this idea, McCoy and Platt (2005) used a visual gambling task to assess whether previously-reported contextual modulation of neuronal activity in CGp reflects subjective value or the objective properties of available rewards. Monkeys were given a choice between two options on a computer monitor. Choosing the safe option always re sulted in a medium-sized squirt of juice. Choosing the risky option resulted in a 50% chance of a large squirt of juice and a 50% chance of a small squirt of juice. Surprisingly, monkeys strongly preferred the risky option when both had the same expected value (Figure 29.9a). In fact, monkeys continued to choose the risky option even when the probability of a larger than average reward was only 1/3. CGp neurons closely mirrored this behavioral bias, rather than representing the objective value of each target (Figure 29.9b). Further, CGp neuronal activity was correlated with subjective target value estimated from the history of monkeys' choices and rewards received (McCoy and Platt, 2005). These data are consistent with the hypothesis that CGp contributes to decision making by evaluating external events and actions with respect to the subjective preferences of the animal. One concern might be whether this modulation of neuronal

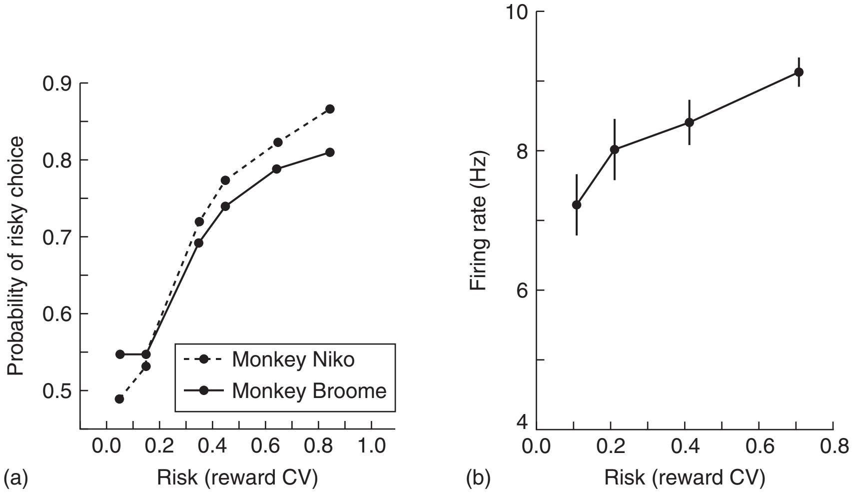


FIGURE 29.9 CGp neurons report the subjective value of a gamble. Monkeys chose between a risky option and a safe option matched for expected value while neuronal activity in posterior cingulate cortex was recorded. (a) Two monkeys preferred the risky option, and preference increased with increasing risk. (b) The firing rate preceding reward delivery increased with increasing risk, thus matching the subjective preferences of the monkeys.

activity in CGp associated with choosing risky options reflects arousal. However, heart rate (a somatic correlate of physiological arousal) did not vary between high-risk and low-risk blocks of trials.

Risk-seeking by monkeys violates expectations based on the assumption of diminishing marginal utility of rewards. One possible explanation for monkeys' preferences could be that monkeys focus on the large reward and ignore bad outcomes. Hayden and Platt tested this hypothesis by examining the relationship between risk preference and delay between trials (Hayden *et al.*, 2007). They found that monkeys' preference for risk declines with increasing delays and reverses when delays increase beyond 45 seconds. These results can be explained by "string theory" (Rachlin, 2000), which proposes that the salience of the large reward, and the expected delay until that reward can be obtained, influence valuation of a risky option. In this model, the value of the risky option is only updated following salient payoffs (in this case, the large rewards) and all rewards are discounted by the delay since the last reward. When the time between trials is short the value of the large reward delivered for choosing the risky option remains essentially unmitigated by the delay, but when the time between trials is long the large reward is heavily discounted. This model accurately predicts the systematic preference reversal of monkeys in this task. Similar processes have been proposed to operate in humans who pursue the immediate, intense "high" of certain drugs of abuse, while simultaneously discounting the delayed, longer-term "low" of withdrawal (Bickel *et al.*, 1999; Bernheim and Rangel, 2004). Alternatively, monkeys may have a concave utility function for reward when the time between trials is short, but the utility function becomes convex when the time between trials is long. In principle, these possibilities might be distinguished using neurophysiological data.

Attention, Learning, and the Posterior Cingulate Cortex

The observation that some CGp neurons respond equivalently to the delivery of larger than average rewards and to the omission of predicted rewards is consistent with a role for this area in highlighting motivationally significant events for the evaluation and control of action (McCoy *et al.*, 2003). The correspondence of neuronal activity in CGp with risk preferences may also reflect subjective focus on the highly salient jackpots occasionally delivered for choosing the risky option (Hayden and Platt, 2007). In addition to guiding future action selection, neuronal activity in CGp may also play a role in learning. Reward modulation of neuronal activity in CGp is consistent with attentional theories of learning, which posit that reward prediction errors highlight motivationally significant events (Pearce and Hall, 1980). According to this idea, neuronal activity correlates with the extent to which outcomes differ from expectations, either positive or negative. While such a signal would not carry information about what needs to be learned, such a signal would instruct when and how effectively learning should occur. Some of the value-related modulations in CGp may thus contribute to attention for learning. Consistent with this idea, lesions of posterior cingulate cortex in rabbits impair certain forms of associative conditioning, and neuronal activity in this same area changes systematically during learning (Gabriel, 1990). However, the homology of posterior cingulate cortex in rabbits to CGp in primates remains uncertain, and thus this hypothesis awaits further study.

NEURO-COGNITIVE MODELS OF CHOICE

As described in previous sections, neuronal correlates of value exist in the OFC, CGp, and LIP. In addition, other studies suggest that value might also be encoded by neurons in numerous other brain regions, including dorsolateral prefrontal cortex (Lee and Wang, 2008), premotor cortex (Roesch and Olson, 2003), frontal eye-fields (Roesch and Olson, 2003), supplementary eye-fields (Stuphorn *et al.*, 2000; Roesch and Olson, 2003), superior colliculus (Ikeda and Hikosaka, 2003), basal ganglia (Kawagoe *et al.*, 1998; Doya and Kimura, 2008), amygdale (Paton *et al.*, 2006), and centromedian nucleus of the thalamus (Minamimoto *et al.*, 2005). Although redundancy is hardly an exception in the nervous system[2], – it is reasonable to ask why there are so many representations of value in the primate brain. Here we propose that value signals expressed by different neuronal populations contribute to different mental processes. In sensory areas, value signals may contribute to perceptual attention (a process of choice between different sensory stimuli); in frontal areas, value signals may contribute to economic choice (a choice between different goods); in motor areas, value signals may contribute to action selection (a choice between different motor acts). To discuss this proposal, we shall specifically discuss value representations in OFC and LIP, because neuronal value signals in these areas have been studied in the greatest detail.

Value Modulations in the Sensory, Goods, and Motor Domains

The rationale of our proposal rests on the observation that value representations in different brain areas are not identical. As described in previous sections, neurons in OFC represent the value of goods *per se*, independently of how goods are visually presented to the monkey, and independently of the action with which the monkey reveals its choice (Figure 29.2d). In contrast, neurons in LIP represent value as a modulation of responses encoding the spatial location of the visual stimulus or the eye movement necessary to express the choice. Another important difference is that the value representation in OFC is menu invariant, whereas the value representation in LIP is menu dependent.

In sensory areas, value modulations may be largely coextensive with the allocation of attention. For example, a rich literature shows that attention modulates visual responses in LIP, and this area can be thought of as representing a map of salient locations in the visual field (Goldberg *et al.*, 2002). From a psychological point of view, in most experimental paradigms attention and value are naturally linked (Maunsell, 2004). On the one hand, attention is typically drawn to a particular visual stimulus by selectively increasing the value associated with that stimulus. On the other hand, any manipulation of the value associated with a particular visual stimulus inevitably influences the attention allocated to that stimulus. Thus, attention and value are often confounded (Maunsell, 2004). Several observations are relevant for evaluating the contributions of attention and value to neuronal activity in LIP. First, Bendiksby and Platt demonstrated that both attention and value contribute independently to neuronal activity in LIP when these two factors are dissociated behaviorally. In their study, value and attention modulations jointly determined changes in behavioral performance typically associated with attention, and were independent of the movement made by the animal (Bendiksby and Platt, 2006). Second, value modulations in LIP take the form of a value weight: the response to one particular stimulus is modulated by the ratio of the value of that stimulus to the value sum of all other visual stimuli (Doris and Glimcher, 2004) In other words, the "total value modulation" (i.e., the value modulation integrated over all spatial locations) is a constant and does not change over time. What does change over time, from one moment to the next, is how the total value modulation is distributed across spatial locations. This observation is consistent with an attentional modulation. Indeed, attention is generally conceptualized as a scarce resource that can be competitively allocated to different stimuli or spatial locations[3] (Bundesen, 1990; Desimone and Duncan, 1995), an idea consistent with

---

[2] For example, in the primate brain, arm movements are represented in some six to eight different areas, and eye movements are represented in at least six different areas.

[3] Bundesen, 1990, in fact defines attention as a choice process taking place through value weights.

value modulating perceptual processing and orienting behavior. In conclusion, value signals in sensory areas might underlie perceptual attention.

In motor areas, value signals could contribute to action selection. From a computational point of view, planning and controlling even a simple voluntary motor act (such as an arm reaching movement) is very challenging. It involves resolving multiple problems with infinite solutions and many degrees of freedom (Mussa-Ivaldi and Bizzi, 2000; Wolpert and Ghahramani, 2000). One reasonable hypothesis is that the nervous system might partly solve this problem through mechanisms in which different motor plans compete with one another. For example, in the model of Glimcher and colleagues, a value is associated with each motor plan, and a single motor act eventually emerges through a winner-takes-all process (Glimcher *et al*., 2005). Behavioral results in humans and other species support this view (see Chapter 8 of this volume) Value modulations that exist in genuinely motor areas (Stuphorn *et al*., 2000; Ikeda and Hikosaka, 2003; Roesch and Olson, 2003; Chapter 26 of this volume) might thus contribute to action selection.

Finally, in areas that encode the value of goods *per se*, such as OFC, value signals could underlie economic choice. Consider again a person sitting in a restaurant and choosing between tuna tartare and fried calamari. Presumably, the person assigns a value to the tuna and a value to the calamari, and then compares the two values to make a decision. OFC neurons encoding the *offer value* seem to capture this valuation process, because different neurons encode the value of different available goods. The actual decision might then emerge through a winner-takes-all mechanism, similar to the one thought to mediate action selection.

"Goods-based" and "Actions-based" Models of Economic Choice

The proposal that value signals that are independent of the sensory-motor contingencies of choice (as in OFC) *contribute* to the choice between goods does not necessarily imply that economic choices completely unfold within these neuronal representations of value. With respect to this fundamental issue, two alternative hypotheses can be considered: the "goods-based" model and the "actions-based" model.

According to the goods-based model, economic choice is an independent cognitive module (Fodor, 1983; Pinker, 1997) – a set of mental and neuronal processes that take place in a space where goods are represented as such. This proposal thus assumes a level of mental abstraction – the space of goods – computationally removed from sensory and motor representations. Values are assigned online to the available goods, based on their properties and on the internal state of the animal at the time of choice. Formally, the value assigned to a given good is described as a function V(a_1, a_2, ...) of multiple arguments a_1, a_2... that describe the good, the motivational state of the animal, the choice contingencies (probabilities, costs, delays, etc.), and the behavioral context of choice. The rules and mechanisms that underlie the computation of the neuronal value function are currently unknown, although neuronal responses recorded in OFC demonstrate that this abstract value function *is* computed. The key feature of the goods-based model is that economic choice fully takes place in the space of goods. In other words, when an individual chooses between goods X and Y, values are assigned to the two goods and a decision is made between these values. Once one good is chosen, the individual plans and executes a suitable motor action to implement the choice. According to the good-based model, however, action selection is a process distinguished from and following economic choice. In this sense, the good-based model of economic choice is modular and sequential.

In contrast, according to the actions-based model, economic choices are embedded in premotor processes of action selection. Several models of decision making can be viewed as variations on the actions-based model (see Chapters 26 and 31 of this volume; Dickinson and Balleine, 2002; Glimcher *et al*., 2005). In one version of the model, originally proposed by Skinner (1953), the behavior of the animal is described simply in terms of sensory stimuli and motor responses, and the problem of choice is essentially reduced to a problem of associative learning. During training, the animal learns the association between a set of stimuli and the rewarded motor response; the animal thus develops a look-up table to be consulted in future choices. In a more recent version of the action-based model, proposed by Glimcher and colleagues (2005), values are learned through experience (possibly in an abstract representation) through mechanisms of reinforcement learning (Sutton and Barto, 1998). At the time of choice, values are retrieved and funneled through the action-selection system, such that a value is attached to each possible course of action. Economic choice thus unfolds as a process of action selection, through a winner-takes-all mechanism. In this view, brain areas and neuronal populations responsible for action selection (such as LIP) represent a common pathway for different types of decision making; they are the substrate upon which choices are actually generated (Glimcher *et al*., 2005). Hence, according to the actions-based model, economic choice is fundamentally choice between actions.

Comparing Goods-based and Actions-based Models

A priori, the goods-based and the actions-based models of choice are both legitimate proposals, and we shall now discuss some of their merits.

The actions-based model has traditionally been more prominent. Two lines of reasoning have been presented to support this model. First, because it builds more or less directly on theories of reinforcement learning, the actions-based model is often presented as one aspect of a more general psychological model of behavior. In principle, a unitary model might seem more parsimonious and thus desirable. At the same time, it can be noted that learning and choice are conceptually distinct and dissociable mental processes. For example, there can be choice in the absence of measurable learning (as in the experiments of Padoa-Schioppa and Assad), as there can be learning in the absence of choice (as in classical and instrumental conditioning). Hence, one particular model of learning does not imply one particular mechanism supporting economic choice. Another argument often presented to support actions-based models is that, at the neuronal level, value signals were first observed in brain areas involved in sensory-motor processes (Kawagoe *et al.*, 1998; Platt and Glimcher, 1999; Ikeda and Hikosake, 2003; Roesch and Olson, 2003; Chapter 26 of this volume). However, this is not an argument against the goods-based model, because value signals in sensory-motor areas could underlie action selection but at the same time make no contribution to economic choice.

Two arguments have been presented to support the goods-based model (Padoa-Schioppa and Assad, 2008). First, the goods-based model seems in principle more efficient because it is modular (Simon, 1962; Pinker, 1997). Planning and controlling movements is computationally challenging. In a modular architecture, the nervous system would break down the complex operation [choosing and moving] into two separate and simpler operations, [choosing] and [moving]. If convolving the control of any motor act with economic choice (actions-based model) requires extra computational power, however minimal, the great redundancy of the motor systems would make a non-modular design enormously more expensive than a modular design. Importantly, natural selection does not always find optimal solutions. It is possible that primates may have evolved to make actions-based economic choices, even though this process may be less efficient than making goods-based choices. In this sense, the OFC results reviewed above can be viewed as an existence proof. Indeed, neurons in OFC encoding the value of offered and chosen goods encode the variables necessary for efficient, goods-based economic choices. One weakness of this argument, however, is that while OFC provides an abstract representation of value, there is no evidence yet demonstrating that the entire choice process (including value comparisons) fully takes places in the abstract representation of goods. The key tenet of the good-based model thus remains to be tested.

The other argument presented to favor the goods-based model as opposed to the actions-based models builds on observations in neuropsychology. Actions-based models of choice instantiated in specific sensory-motor areas, such as LIP (Glimcher *et al.*, 2005) or the basal ganglia (see Chapter 26), often overlook the fact that lesions to these areas do not typically influence economic choice behavior *per se*. For example, lesions of the parietal cortex result in visuo-spatial deficits such as hemi-neglect and Balint's syndrome (Colby and Olson, 1999). In contrast, economic choices are typically disrupted by OFC lesions (Bechara *et al.*, 1996; Pasquier and Petit, 1997; Rahman *et al.*, 1999; Fellows and Farah, 2007; Koenigs and Tranel, 2007). In other words, unlike OFC, sensory-motor areas are not strictly necessary for making economic choices, although they are required for implementing them. Notably, while this line of reasoning argues against specific versions of the actions-based model, it does not prove (or explicitly argue for) the goods-based model.

Two important points should be emphasized. First, the goods-based model only applies to choices between goods (economic choices). However, during the normal course of behavior, different valuation processes occur simultaneously as animals make choices in the sensory, goods, and motor domains. Second, behavior often evolves from choice to habit, thus requiring less deliberation – a process mirrored by changes in neuronal circuitry (Graybiel, 2005).

How can the two models be tested more directly? Both models conceptualize economic choice as a two-stage mental process where values are initially assigned to the available goods and a decision (i.e., a comparison between values) is subsequently made (Glimcher *et al.*, 2005; Padoa-Schioppa *et al.*, 2006). Apart from the role of learning, the two models differ on one critical point: according to the goods-based model, choice should be *completely* processed within an abstract representation of goods. Because an abstract representation of value exists in the OFC, distinguishing between the two models requires establishing whether the decision process that follows valuation (that is, the winner-takes-all mechanisms through which different values are compared) takes place in the space of goods or in the space of actions. In principle, this question can be addressed by separating in time the choice between goods and the selection of action. This issue thus remains an important question for future work.

CONCLUSION

Multiple representations of value exist in the primate brain. Specifically, neurons in the orbitofrontal cortex encode the value subjects assign to different goods, independently of how the goods are presented or the action necessary to implement choice. Moreover, the representation of value in the OFC is menu invariant – neuronal responses encoding the value of one particular good do not depend on what other goods are available at the same time. In contrast, neurons in parietal cortex are sensitive to the value of a particular location in space or a particular action. Whereas representations of value in OFC are "absolute," value signals in parietal cortex are "relative" – each neuron is modulated by the ratio of the value of the corresponding location/action to the value sum of all locations/actions. Finally, neurons in posterior cingulate cortex appear to encode an ongoing estimate of the subjective value or salience of extra-personal events and actions; these signals may update value-related information in brains areas like the parietal cortex.

These observations suggest that different neuronal representations of value in different parts of the brain contribute to distinct computational processes. In sensory areas, value-related modulation in neuronal activity might promote selective processing of high-value stimuli by attention. In frontal cortex, neurons representing abstract value could contribute directly to the selection of one among multiple available goods based on subjective preference. Finally, in sensory-motor areas, value modulations might promote the selection of one particular action among many to achieve a behavioral goal.

References

Amador, N., Schlag-Rey, M., and Schlag, J. (2000). Reward-predicting and reward-detecting neuronal activity in the primate supplementary eye field. *J. Neurophysiol.* 84, 2166–2170.

Andersen, R.A., Asanuma, C., Essick, G., and Siegel, R.M. (1990). Corticocortical connections of anatomically and physiologically defined subdivisions within the inferior parietal lobule. *J. Comp. Neurol.* 296, 65–113.

Arana, F.S., Parkinson, J.A., Hinton, E. *et al*. (2003). Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection. *J. Neurosci.* 23, 9632–9638.

Arnaud, A. and Nichole, P. (1982 (1662)). *The Art of Thinking: Port-Royal Logic*. Indianapolis, IN: Bobbs-Merrill.

Balan, P.F. and Gottlieb, J. (2006). Integration of exogenous input into a dynamic salience map revealed by perturbing attention. *J. Neurosci.* 26, 9239–9249.

Baleydier, C. and Mauguiere, F. (1980). The duality of the cingulate gyrus in monkey. Neuroanatomical study and functional hypothesis. *Brain* 103, 525–554.

Barbas, H. and Mesulam, M.M. (1981). Organization of afferent input to subdivisions of area 8 in the rhesus monkey. *J. Comp. Neurol.* 200, 407–431.

Barbas, H. and Mesulam, M.M. (1985). Cortical afferent input to the principalis region of the rhesus monkey. *Neuroscience* 15, 619–637.

Barbas, H. and Pandya, D.N. (1989). Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.* 286, 353–375.

Basso, M.A. and Wurtz, R.H. (1997). Modulation of neuronal activity by target uncertainty. *Nature* 389, 66–69.

Bechara, A., Tranel, D., Damasio, H. and Damasio, A.R. Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex,* 6, 215–225.

Bendiksby, M.S. and Platt, M.L. (2006). Neural correlates of reward and attention in macaque area LIP. *Neuropsychologia* 44, 2411–2420.

Berman, R.A., Colby, C.L., Genovese, R. *et al*. (1999). Cortical networks subserving pursuit and saccadic eye movements in humans: an FMRI study. *Human Brain Mapp.* 8, 125–209.

Bernheim, B.D. and Rangel, A. (2004). Addiction and cue-triggered decision processes. *Am. Econ. Rev.* 94, 1558–1590.

Bickel, W.K., Odum, A.L., and Madden, G.J. (1999). Impulsivity and cigarette smoking: delay discounting in current, never, and ex-smokers. *Psychopharmacology* 146, 447–454.

Blatt, G.J., Andersen, R.A., and Stoner, G.R. (1990). Visual receptive field organization and cortico-cortical connections of the lateral intraparietal area (area LIP) in the macaque. *J. Comp. Neurol.* 299, 421–445.

Braak, H., Braak, E., Yilmazer, D. *et al*. (1996). Pattern of brain destruction in Parkinson's and Alzheimer's diseases. *J. Neural Transm.* 103, 455–490.

Bundesen, C. (1990). A theory of visual attention. *Psychological Rev.* 97, 523–547.

Cammalleri, R., Gangitano, M., D'Amelio, M. *et al*. (1996). Transient topographical amnesia and cingulate cortex damage: a case report. *Neuropsychologia* 34, 321–326.

Carmichael, S.T. and Price, J.L. (1994). Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *J. Comp. Neurol.* 346, 366–402.

Carmichael, S.T. and Price, J.L. (1995). Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *J. Comp. Neurol.* 363, 642–664.

Cavada, C. and Goldman-Rakic, P.S. (1989a). Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *J. Comp. Neurol.* 287, 422–445.

Cavada, C. and Goldman-Rakic, P.S. (1989b). Posterior parietal cortex in rhesus monkey: I. Parcellation of areas based on distinctive limbic and sensory corticocortical connections. *J. Comp. Neurol.* 287, 393–421.

Cavada, C., Company, T., Tejedor, J. *et al*. (2000). The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cerebral Cortex* 10, 220–242.

Cheney, D.L. and Seyfarth, R.M. (1990). *How Monkeys See the World: Inside the Mind of Another Species*. Chicago, IL: Chicago University Press, x 377.

Choi, S., Fisman, R., Gale, D., and Kariv, S. (2008). Consistency and heterogeneity of individual behavior under uncertainty. *Am. Econ. Rev.* 97, 1858–1876.

Coe, B., Tomihara, K., Matsuzawa, M., and Hikosaka, O. Visual and anticipatory bias in three cortical eye fields of the monkey during an adaptive decision-making task. *J. Neurosci*., 22, 5081–5090.

Colby, C.L. and Olson, C.R. (1999). Spatial cognition. In: M.J. Zigmond, F.E. Bloom, S.C. Landis. *et al*. (eds), *Fundamental Neuroscience*. San Diego, CA: Academic Press, pp. 1363–1383.

Colby, C.L., Duhamel, J.R., and Goldberg, M.E. (1996). Visual, presaccadic, and cognitive acivation of single neurons in monkey lateral intraparietal area. *J. Neurophysiol.* 76, 2841–2852.

Damasio, H., Grabowski, T., Frank, R. *et al*. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science* 264, 1102–1105.

Dean, H.L. and Platt, M.L. (2006). Allocentric spatial referencing of neuronal activity in macaque posterior cingulate cortex. *J. Neurosci.* 26, 1117–1127.

Dean, H.L., Crowley, J.C., and Platt, M.L. (2004). Visual and saccade-related activity in macaque posterior cingulate cortex. *J. Neurophysiol.* 92, 3056–3068.

Deaner, R.O., Khera, A.V., and Platt, M.L. (2005). Monkeys pay per view: adaptive valuation of social images by rhesus macaques. *Curr. Biol.* 15, 543–548.

Desimone, R. and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222.

Dickhaut, J., McCabe, K., Nagode, J.C. *et al*. (2003). The impact of the certainty context on the process of choice. *Proc. Natl Acad. Sci. USA* 100, 3536–3541.

Dickinson, A. and Balleine, B.W. (2002). The role of learning in the operation of motivational systems. In: C.R. Gallistel (ed.), *Learning, Motivation and Emotion. Steven's Handbook of experimental Psychology*, Vol. 3. New York, NY: John Wiley & Sons, pp. 497–533.

Dorris, M.C. and Glimcher, P.W. (2004). Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44, 365–378.

Fellows, L.K. and Farah, M.J. (2007). The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cerebral Cortex* 17, 2669–2674.

Feierstein, C.E., Quirk, M.C., Uchida, N. *et al*. (2006). Representation of spatial goals in rat orbitofrontal cortex. *Neuron* 51, 495–507.

Fletcher, W.A. (1994). Ophthalmological aspects of Alzheimer's disease. *Curr. Opin. Ophthalmol.* 5, 38–44.

Flitman, S., O'Grady, J., Cooper, V., and Grafman, J. (1997). PET imaging of maze processing. *Neuropsychologia* 35, 409–420.

Fodor, J.A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.

Gabriel, M. (1990). Functions of anterior and posterior cingulate cortex during avoidance learning in rabbits. *Prog. Brain Res.* 85, 467–482, discussion 482–483.

Gabriel, M., Vogt, B.A., Kubota, Y. *et al*. (1991). Training-stage related neuronal plasticity in limbic thalamus and cingulate cortex during learning: a possible key to mnemonic retrieval. *Behav. Brain Res.* 46, 175–185.

Ghaem, O., Mellet, E., Crivello, F. *et al*. (1997). Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. *NeuroReport* 8, 739–744.

Glimcher, P.W., Dorris, M.C., and Bayer, H.M. (2005). Physiological utility theory and the neuroeconomics of choice. *Games Econ. Behav.* 52, 213–256.

Gnadt, J.W. and Andersen, R.A. (1988). Memory related motor planning activity in posterior parietal cortex of macaque. *Exp. Brain Res.* 70, 216–220.

Gold, J.I. and Shadlen, M.N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* 5, 10–16.

Goldberg, M.E., Colby, C.L., and Duhamel, J.R. (1990). Representation of visuomotor space in the parietal lobe of the monkey. *Cold Spring Harbour Symp. Quant. Biol.* 55, 729–739.

Goldberg, M.E., Bisley, J., Powell, K.D. *et al*. (2002). The role of the lateral intraparietal area of the monkey in the generation of saccades and visuospatial attention. *Ann. N Y Acad. Sci.* 956, 205–215.

Goodall, E.B. and Carey, R.J. (1975). Effects of d- versus l-amphetamine, food deprivation, and current intensity on self-stimulation of the lateral hypothalamus, substantia nigra, and medial frontal cortex of the rat. *J. Comp. Physiol. Psychol.* 89, 1029–1045.

Gottlieb, J.P., Kusunoki, M., and Goldberg, M.E. (1998). The representation of visual salience in monkey parietal cortex. *Nature* 391, 481–484.

Grace, R.C. (1993). Violations of transitivity: implications for a theory of contextual choice. *J. Exp. Anal. Behav.* 60, 185–201.

Graybiel, A.M. (2005). The basal ganglia: learning new tricks and loving it. *Curr. Opin. Neurobiol.* 15, 638–644.

Hrdy, S.B. and Whitten, P.L. (1987). Patterning of sexual activity. In: B.B. Smuts, R.M. Seyfarth, R.W. Wrangham, and T.T. Struhsake (eds), *Primate Societies*. Chicago, IL: University of Chicago Press, pp. 370–384.

Hayden, B.Y. and Platt, M.L. (2007). Temporal discounting predicts risk sensitivity in rhesus macaques. *Curr. Biol.* 17, 49–53.

Hayden, B.Y., Parikh, P.C., Deaner, R.O., and Platt, M.L. (2007). Economic principles motivating social attention in humans. *Proc. Biol. Sci.* 274, 1751–1756.

Heeger, D.J. (1993). Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. *J. Neurophysiol.* 70, 1885–1898.

Hikosaka, O., Nakamura, K., and Nakahara, H. (2006). Basal ganglia orient eyes to reward. *J. Neurophysiol.* 95, 567–584.

Hof, P.R., Archin, N., Osmand, A.P. *et al*. (1993). Posterior cortical atrophy in Alzheimer's disease: analysis of a new case and re-evaluation of a historical report. *Acta Neuropathol.* 86, 215–223.

Hof, P.R., Vogt, B.A., Bouras, C., and Morrison, J.H. (1997). Atypical form of Alzheimer's disease with prominent posterior cortical atrophy: a review of lesion distribution and circuit disconnection in cortical visual pathways. *Vision Res.* 37, 3609–3625.

Hopfinger, J.B., Buonocore, M.H., and Mangun, G.R. (2000). The neural mechanisms of top-down attentional control. *Nat. Neurosci.* 3, 284–291.

Hopfinger, J.B., Woldorff, M.G., Fletcher, E.M., and Mangun, G.R. (2001). Dissociating top-down attentional control from selective perception and action. *Neuropsychologia* 39, 1277–1291.

Horvitz, J.C. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non–reward events. *Neuroscience* 96, 651–656.

Huettel, S.A., Stowe, C.J., Gordon, E.M. *et al*. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron* 49, 765–775.

Hsu, M., Bhatt, M., Adolphs, R. *et al*. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310, 1680–1683.

Ikeda, T. and Hikosaka, O. (2003). Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron* 39, 693–700.

Ito, S., Stuphorn, V., Brown, J.W., and Schall, J.D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science* 302, 120–122.

Kable, J.W. and Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.

Kalia, M. and Whitteridge, D. (1973). The visual areas in the splenial sulcus of the cat. *J. Physiol.* 232, 275–283.

Kandel, E.R., Schwartz, J.H., and Jessell, T.M. (eds) (2000). *Principles of Neural Science*, Vol. XLI. New York, NY: McGraw-Hill.

Katayama, K., Takahashi, N., Ogawara, K., and Hattori, T. (1999). Pure topographical disorientation due to right posterior cingulate lesion. *Cortex* 35, 279–282.

Kawagoe, R., Takikawa, Y., and Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* 1, 411–416.

Kim, Y.H., Gitelman, D.R., Nobre, A.C. *et al*. (1999). The large-scale neural network for spatial attention displays multifunctional overlap but differential asymmetry. *NeuroImage* 9, 269–277.

Koenigs, M. and Tranel, D. (2007). Irrational economic decision-making after ventromedial prefrontal damage: evidence from the Ultimatum Game. *J. Neurosci.* 27, 951–956.

Klein, J.T., Deaner, R.O., and Platt, M.L. (2008). Neural correlates of social target value in macaque parietal cortex. *Curr. Biol.* 18, 419–424.

Kreps, D.M. (1990). *A Course in Microeconomic Theory*. Princeton, NJ: Princeton University Press.

Kuhnen, C.M. and Knutson, B. (2005). The neural basis of financial risk taking. *Neuron* 47, 763–770.

Leon, M.I. and Shadlen, M.N. (1999). Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* 24, 415–425.

Maddock, R.J. (1999). The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. *Trends Neurosci.* 22, 310–316.

Maddock, R.J., Garrett, A.S., and Buonocore, M.H. (2003). Posterior cingulate cortex activation by emotional words: fMRI evidence from a valence decision task. *Human Brain Mapp.* 18, 30–41.

Maratos, E.J., Dolan, R.J., Morris, J.S. *et al*. (2001). Neural activity associated with episodic memory for emotional context. *Neuropsychologia* 39, 910–920.

Maunsell, J.H. (2004). Neuronal representations of cognitive state: reward or attention? *Trends Cogn. Sci.* 8, 261–265.

Mazur, J.E. and Coe, D. (1987). Tests of transitivity in choices between fixed and variable reinforcer delays. *J. Exp. Anal. Behav.* 47, 287–297.

McCoy, A.N. and Platt, M.L. (2005). Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat. Neurosci.* 8, 1220–1227.

McCoy, A.N., Crowley, J.C., Haghighian, G. *et al*. (2003). Reward signals in posterior cingulate cortex. *Neuron* 40, 1031–1040.

Mesulam, M.M., Nobre, A.C., Kim, Y.H. *et al*. (2001). Heterogeneity of cingulate contributions to spatial attention. *Neuroimage* 13, 1065–1072.

Minamimoto, T., Hori, Y., and Kimura, M. (2005). Complementary process to response bias in the centromedian nucleus of the thalamus. *Science* 308, 1798–1801.

Minoshima, S., Giordani, B., Barent, S. *et al*. (1997). Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann. Neurol.* 42, 85–94.

Morecraft, R.J., Geula, C., and Mesulam, M.M. (1993). Architecture of connectivity within a cingulo-fronto-parietal neurocognitive network for directed attention. *Arch. Neurol.* 50, 279–284.

Muller-Preuss, P. and Jurgens, U. (1976). Projections from the "cingular" vocalization area in the squirrel monkey. *Brain Res.* 103, 29–43.

Mussa-Ivaldi, F.A. and Bizzi, E. (2000). Motor learning through the combination of primitives. *Phil. Trans. R. Soc. Lond. B*, 355, 1755–1769.

Navarick, D.J. and Fantino, E. (1972). Transitivity as a property of choice. *J. Exp. Anal. Behav.* 18, 389–401.

Niki, H. and Watanabe, M. (1979). Prefrontal and cingulate unit activity during timing behavior in the monkey. *Brain Res.* 171, 213–224.

O'Doherty, J.P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* 14, 769–776.

Olson, C.R. and Musil, S.Y. (1992). Posterior cingulate cortex: sensory and oculomotor properties of single neurons in behaving cat. *Cerebral Cortex* 2, 485–502.

Olson, C.R., Musil, S.Y., and Goldberg, M.E. (1996). Single neurons in posterior cingulate cortex of behaving macaque: eye movement signals. *J. Neurophysiol.* 76, 3285–3300.

Padoa-Schioppa, C. (2007). Orbitofrontal cortex and the computation of economic value. *Ann. NY Acad. Sci.* 1121, 232–253.

Padoa–Schioppa, C. and Assad, J.A. (2006). Neurons in orbitofrontal cortex encode economic value. *Nature* 441, 223–226.

Padoa-Schioppa, C. and Assad, J.A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat. Neurosci.* 11, 95–102.

Padoa-Schioppa, C., Jandolo, L., and Visalberghi, E. (2006). Multistage mental process for economic choice in capuchins. *Cognition* 99, B1–b13.

Pandya, D.N., van Hoesen, G.W., and Mesulam, M.M. (1981). Efferent connections of the cingulate gyrus in the rhesus monkey. *Exp. Brain Res.* 42, 319–330.

Pasquier, F. and Petit, H. (1997). Frontotemporal dementia: its rediscovery. *Eur. Neurol.* 38, 1–6.

Paton, J.J., Belova, M.A., Morrison, S.E., and Salzman, C.D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439, 865–870.

Pearce, J.M. and Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Rev.* 87, 532–552.

Pine, D.S., Grun, J., Maguire, E.A. *et al*. (2002). Neurodevelopmental aspects of spatial navigation: a virtual reality fMRI study. *NeuroImage* 15, 396–406.

Pinker, S. (1997). *How the Mind Works*, Vol. XII. New York, NY: Norton.

Platt, M.L. and Glimcher, P.W. (1997). Responses of intraparietal neurons to saccadic targets and visual distractors. *J. Neurophysiol.* 78, 1574–1589.

Platt, M.L. and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.

Powell, E.W. (1978). The cingulate bridge between allocortex, isocortex and thalamus. *Anat. Rec.* 190, 783–793.

Price, J.L. (2007). Definition of the orbital cortex in relation to specific connections with limbic and visceral structures, and other cortical regions. *Ann NY Acad. Sci.* 1121, 54–71.

Rachlin, H. (2000). *The Science of Self–control*. Cambridge, MA: Harvard University Press.

Rahman, S., Sahakian, B.J., Hodges, J.R. *et al*. (1999). Specific cognitive deficits in mild frontal variant of frontotemporal dementia. *Brain* 122, 1469–1493.

Roesch, M.R. and Olson, C.R. (2003). Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *J. Neurophysiol.* 90, 1766–1789.

Roesch, M.R. and Olson, C.R. (2005). Neuronal activity in primate orbitofrontal cortex reflects the value of time. *J. Neurophysiol.* 94, 2457–2471.

Roesch, M.R., Taylor, A.R., and Schoenbaum, G. (2006). Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* 51, 509–520.

Roitman, J.D., Brannon, E.M., and Platt, M.L. (2007). Monotonic coding of numerosity in macaque lateral intraparietal area. *PLoS Biol* 5, e208.

Rolls, E.T., Sienkiewicz, Z.J., and Yaxley, S. (1989). Hunger modulates the responses to gustatory stimuli of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey. *Eur. J. Neurosci.* 1, 53–60.

Salzman, C.D., Belova, M.A., and Paton, J.J. (2005). Beetles, boxes and brain cells: neural mechanisms underlying valuation and learning. *Curr. Opin. Neurobiol.* 15, 721–729.

Sato, M. and Hikosaka, O. (2002). Role of primate substantia nigra pars reticulate in reward-oriented saccadic eye movement. *J. Neurosci.* 22, 2363–2373.

Schwartz, O. and Simoncelli, E.P. (2001). Natural signal statistics and sensory gain control. *Nat. Neurosci.* 4, 819–825.

Selemon, L.D. and Goldman–Rakic, P.S. (1988). Common cortical and subcortical targets of the dorsolateral prefrontal and posterior parietal cortices in the rhesus monkey: evidence for a distributed neural network subserving spatially guided behavior. *J. Neurosci.* 8, 4049–4068.

Shafir, S. (1994). Intransitivity of preferences in honey bees: support for "comparative" evaluation of foraging options. *Animal Behav.* 48, 55–67.

Shidara, M. and Richmond, B.J. (2002). Anterior cingulate: single neuronal signals related to degree of reward expectancy. *Science* 296, 1709–1711.

Sikes, R.W. and Vogt, B.A. (1992). Nociceptive neurons in area 24 of rabbit cingulate cortex. *J. Neurophysiol.* 68, 1720–1732.

Simon, H.A. (1962). The architecture of complexity. *Proc. Am. Phil. Soc.* 106, 467–482.

Skinner, B.F. (1953). *Science and Human Behavior*. New York, NY: Macmillan.

Small, D.M., Zatorre, R.J., Dagher, A. *et al*. (2001). Changes in brain activity related to eating chocolate: from pleasure to aversion. *Brain* 124, 1720–1733.

Small, D.M., Gitelman, D.R., Gregory, M.D. *et al*. (2003). The posterior cingulate and medial prefrontal cortex mediate the anticipatory allocation of spatial attention. *NeuroImage* 18, 633–641.

Spence, S.J., Silverman, J.A., and Corbett, D. (1985). Cortical and ventral tegmental systems exert opposing influences on self-stimulation from the prefrontal cortex. *Behav. Brain Res.* 17, 117–124.

Stuphorn, V., Taylor, T.L., and Schall, J.D. (2000). Performance monitoring by the supplementary eye field. *Nature* 408, 857–860.

Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.

Sutton, R.S. and Barto, A.G. (1998). Reinforcement Learning: An Introduction, Vol. XVIII. Cambridge, MA: MIT Press.

Takahashi, N., Kawamura, M., Shiota, J. *et al*. (1997). Pure topographic disorientation due to right retrosplenial lesion. *Neurology* 49, 464–469.

Thorpe, S.J., Rolls, E.T., and Maddison, S. (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp. Brain Res.* 49, 93–115.

Tremblay, L. and Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature* 398, 704–708.

Tversky, A. (1969). The intransitivity of preferences. *Psychological Rev.* 76, 31–48.

Vogt, B.A. and Pandya, D.N. (1987). Cingulate cortex of the rhesus monkey: II. Cortical afferents. *J. Comp. Neurol.* 262, 271–289.

Wallis, J.D. (2007). Orbitofrontal cortex and its contribution to decision-making. *Annu. Rev. Neurosci.* 30, 31–56.

Watanabe, K., Lauwereyns, J., and Hikosaka, O. (2003). Neural correlates of rewarded and unrewarded eye movements in the primate caudate nucleus. *J. Neurosci.* 23, 10052–10057.

Wolpert, D.M. and Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nat. Neurosci.* 3(Suppl.), 1212–1217.

Yamasaki, H., LaBar, K.S., and McCarthy, G. (2002). Dissociable prefrontal brain systems for attention and emotion. *Proc. Natl Acad. Sci. USA* 99, 11447–11451.

Yeterian, E.H. and van Hoesen, G.W. (1978). Cortico-striate projections in the rhesus monkey: the organization of certain cortico-caudate connections. *Brain Res.* 139, 43–63.

CHAPTER

30

The Trouble with Choice: Studying Decision Variables in the Brain

Greg S. Corrado, Leo P. Sugrue, Julian R. Brown and William T. Newsome

OUTLINE

INTRODUCTION

Our lives are defined by the decisions we make. From simple acts like picking an item from a lunch menu, to complex choices such as selecting a partner or judging someone guilty of a crime, our decisions vary hugely in complexity and consequence. Given the myriad external and internal factors that influence our choices, it might seem surprising that the underlying machinery of decision making is amenable to scientific analysis at all. Yet in recent years psychologists, economists, and neuroscientists have made significant inroads into this intriguing but mysterious domain. The inquiry has advanced farthest in the arena of sensory-based decision making, where subjects are asked to judge an objective quality of an external stimulus (reviewed in Romo and Salinas, 2003, and Gold and Shadlen, 2007). In this chapter, however, we will focus on the more elusive area of "value-based" decision making, which is a topic of primary interest in neuroeconomics (Sugrue *et al.*, 2005).

It is useful to start with three definitions. First, for our purposes, a *decision* occurs when an organism, confronted by several discrete options, evaluates the merits of each and selects one to pursue. In the laboratory, we study decisions by placing subjects in a controlled environment and offering them two or more mutually exclusive options. Subjects communicate

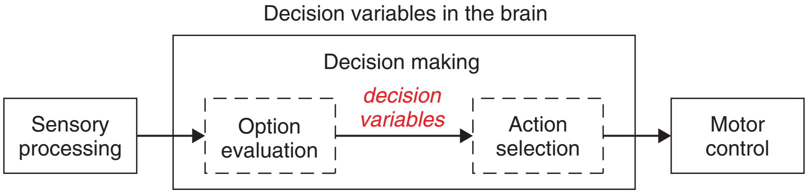


FIGURE 30.1 Decision variables in the brain. In an (over) simplified feed-forward model of the brain, sensory systems provide data to decision-making circuits, which in turn direct motor systems. The machinery of decision making can be further subdivided into mechanisms that evaluate options and those that subsequently select actions. Decision variables can then be operationally defined as the currency of communication between these two sub-stages.

their choices by making overt motor responses such as button presses or eye movements, motivated by the promise of reward (or the threat of punishment) contingent on the responses they make.

Second, a *value-based* decision is one in which the organism's choice is not mandated by the immediate sensory characteristics of the options (as when reading an eye chart at the doctor's office), but rather by the subjective experience and preference of the individual (as when selecting a fine wine in a restaurant). Typically, there are no unambiguously "correct" answers in value-based decisions, which is where the trouble begins from a scientific perspective. Unlike simple sensory judgments, choices rendered in value-based tasks depend on unseen factors internal to the decision maker, and it is here that our third definition comes into play[1].

Decision variables are quantities internal to the subject's decision process that summarize properties of the available behavioral options relevant to guiding choice. As illustrated in Figure 30.1, decision variables can be thought of as linking the processes of option evaluation and action selection. These variables are by their very nature subjective, expressing the decider's private estimation of the attractiveness of the available options. To the outside observer such internal states are necessarily hypothetical – indeed, behaviorists might go so far as to call them fictitious. In contrast, neuroeconomists, guided by the conviction that decisions are mediated by scientifically accessible processes within the brain, embrace the study of internal decision variables as a central goal. We would like to understand where and how these variables are computed within the brain, and how they subsequently guide action selection. Our success or failure in this endeavor will ultimately be judged by our ability to "read out" the state of a decision variable directly from measurements of brain activity, and by our ability to produce predictable changes in choice behavior by manipulating decision variables at the neural level. Before reaching either of these goals, however, neuroeconomists must first confront and solve two related problems.

First, we must find a way to deal with the hypothetical nature of decision variables. We have no *a priori* way of knowing what decision variables guide particular choices. Even in very simple value-based decision-making situations, the factors guiding choices may include any property of a behavioral option that might impact its desirability: reward magnitude, reward quality, risk, uncertainty, delay, cost, etc. While each of these elements might individually constitute a decision variable in its raw form, they might also be transformed or combined in some complicated way (as by the economist's utility function). Moreover, any of these variables – raw, transformed, or combined – might be encoded within the brain in absolute terms, relative to other available options, or even relative to the subject's expectations. In yet more complex situations, decision variables might also include additional hidden factors ranging from the physiological state of the subject (e.g. hunger) to subtle social pressures to make the "right" choice.

The second problem follows directly from the first: because we do not know *a priori* what decision variables are actually used by the brain, the standard correlative methods employed by systems neuroscientists are seriously compromised. Historically, success in systems neuroscience has come from manipulating a variable of interest (i.e., a sensory stimulus or a motor act), comparing neural activity across experimental conditions, and attributing observed differences to functional processing within the central nervous

---

[1] Occasionally, there are sensory judgments that lack an objectively "correct" answer, such as the famed Necker Cube, or value-based decisions that have an obviously "correct" answer under only very weak assumptions – for example, would you rather have a 50% chance of winning $10 or a 50% chance of losing $10? – but these cases are rare. Typically, sensory judgments hinge on discovering some real objective state of the external world, and value judgments depend on the assessment of outcomes whose worth can be genuinely contentious and idiosyncratic to the individual decider. Of course, most real-world decisions require both sensory and value-based assessments: what is the state of the world, and how desirable is that state?

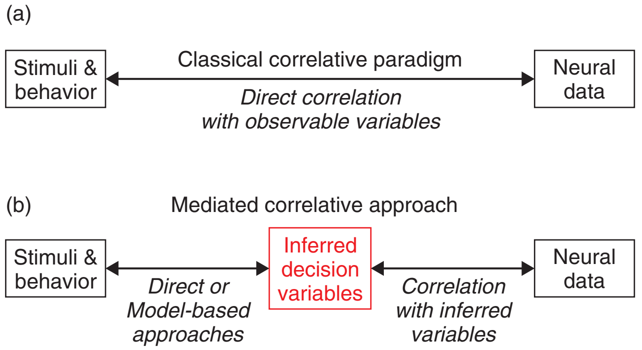


FIGURE 30.2 Topology of correlative methods. (a) The classical correlative approach looks for direct correlations between observable variables, such as the presented sensory stimuli or the subject's overt behavior, and physiological signals, such as neural firing rates or blood oxygen-level dependent (BOLD) magnetic resonance signals (fMRI). (b) In the study of decision making, it is generally necessary to use a mediated correlative approach – inferring abstract decision variables that depend on stimuli and predict behavior, and in turn correlating these computed variables with physiological measurements.

system (Figure 30.2a)[2]. Studying decision variables with this approach becomes difficult when we do not already know what decision variables operate in a given situation.

We can try to solve these problems by “bootstrapping” our way into the decision making system. To do this we make a reasonable guess about the decision variables operating in a given task, and then try to identify correlates of these variables in the brain (Figure 30.2b). The neural signals that we identify through this process may help us reject certain candidate decision variables and form new hypotheses about others. However, identifying meaningful signals in the first place hinges on the accuracy of our initial guess at the candidate decision variables. To improve the accuracy of this guess, we must begin with what *is* observable – the subject's choices. Through careful analysis of behavior, we hope to form good approximations to the hidden decision variables.

Because this process is so crucial, we devote this chapter to a review of the most promising approaches to recovering hidden decision variables from behavior. First, we consider “direct” methods of inferring decision variables from patterns of choice. We then consider “model-based” approaches that construct an explicit model of the entire decision-making process, before going on to describe approaches to validating candidate decision variables by rigorously comparing their predictions to actual behavior. Finally, we discuss some recent studies in which techniques for estimating decision variables have been successfully applied to search for neural correlates.

DIRECT METHODS FOR ACCESSING DECISION VARIABLES

Several techniques are available for directly inferring subjective decision variables from behavior. These “model-free” methods are simple in the sense that they do not require mathematical descriptions of how decision variables are computed.[3] However, sometimes this simplicity comes at the cost of cumbersome or fragile experimental design. To begin with, we consider the most obvious approach to accessing decisions variables: self-report.

Direct Self-report

No approach to accessing internal states could be more straightforward than asking subjects to simply report their feelings. For example, to assess how a subject values a particular option, we could present the option and ask the subject to rate its value on a linear scale. Simple though it is, for a number of reasons direct self-report is rarely used as a primary technique for accessing decision variables. Psychologists stress subjects' inability to access or correctly identify the true cause of their actions, and the vulnerability of their answers to the precise manner in which the question is framed (see Ross and Nisbett, 1991). Economists, on the other hand, criticize the “lack of incentive” for subjects to report their beliefs truthfully, in the absence of which subjects might report randomly, or even deliberately deceive. Neurophysiologists, meanwhile, dismiss the approach entirely, citing the almost insuperable difficulties in applying such methods to non-human animals. Regardless of the reason, the consensus is clear: unadorned self-report is too unreliable and limited a method to provide robust metrics of hidden decision

---

[2] The “variable of interest” might be something as simple as the pitch of a presented auditory tone (e.g., Reale and Imig, 1980), or something as subtle as the attribution of emotional states to self versus other (e.g. Ochsner *et al*., 2004). Similarly, the “measure of neural activity” might be any available measure of brain physiology: the delivery of oxygen to a particular brain region measured via fMRI, the firing of action potentials by individual neurons within that area measured through electrophysiology, or even the strength of synaptic connectivity between pair of neurons measured by future technologies.

[3] Though some “model-free” techniques do leverage assumptions about how decision variables guide action selection or how decision variables are combined, these conditions, are more modest than the detailed mathematical models of decision making we consider later in the section titled “Model-based Approaches”.

variables. Fortunately, there are refinements of self-report techniques that promise greater reliability.

Incentive Compatible Mechanism Design

To address the economist's primary criticism of self-report methods for assessing value, we can employ the principles of *incentive compatible mechanism design*[4]. The idea itself is reasonably simple: ask subjects to report their preference for a particular option in units we understand (such as dollars), and give them a reason to answer honestly. The mechanism employed is usually some sort of virtual auction, where the "prize" is the option the subject is valuing and the "bid" is the value the subject reports. By participating in this virtual auction subjects communicate their valuation of options, much like direct self-report. In this case, however, it is incentive compatible for subjects to report their valuation truthfully because the details of the auction ensure that if they lie, they will on average be less happy with the final outcome than if they tell the truth.[5] Such auction-based incentive compatible self-report mechanisms have been used to extract valuations of items from human subjects that predict subjects' choices in other contexts and correlate with activity in neural circuits implicated in the process of valuation (see, for example, Plassmann *et al.*, 2007).

While this approach meets the economists' demands for motivating subjects toward honesty, psychologists and neurophysiologists tend to be less satisfied. The psychologist's chief objection is that underlying decision variables are intrinsically difficult to access via introspection, even under proper motivation. To wit, one might wonder if eBay's auction selling prices reflect the bidder's true cold assessment of the value of an item, or are more often than not contaminated by the thrill of competition or clouded by the abstractions of credit cards and PayPal. For the neurophysiologist, meanwhile, who struggles to study neural activity in experimental animals that lack even basic facility with the abstractions of language or currency, any method that relies upon *direct* reports of decision variables, incentive compatible or not, is of limited practical usefulness.

Simple Revealed Preference

Simple *revealed-preference* techniques offer another alternative to self-report, allowing us to infer an option's relative value directly from patterns of choice across pairs of options[6]. At their core, these revealed-preference methods rely on the relatively conservative assumption that when choosing between two options A and B, a subject will choose A more frequently than B if (and only if) A is more desirable:

$$D_A > D_B \leftrightarrow p(choice = A) > p(choice = B) \qquad (30.1)$$

where we have used D to denote the total subjective desirability of an option[7], presumed to combine all the subject's preferences over possible outcomes, the probability of outcomes, risk, cost, etc.

An example of a simple revealed-preference design is one in which subjects are given a large number of paired choices between n units of reward X, and m units of reward Y. By finding the point at which a subject is indifferent to the two options the experimenter can infer the effective exchange rate between X and Y (for example, three units of X are as desirable as two units of Y). Using such *indifference points*, one can obtain an ordinal ranking of the desirability of any amount of either reward quantity, or even combinations of the two[8]. The ranking of option values can then be used to search for neural correlates, thereby identifying candidate structures or circuits for

---

[4] For those unfamiliar with the term, *mechanism design* is the field dedicated to the design of social institutions such as voting systems, auction rules, simple games, and market structures. *Incentive compatibility*, meanwhile is the lofty objective that these institutions should function as the designer intended even if the agents who make up the social group are perfectly selfish and fiendishly clever – i.e. that the "incentives" of the game should render individual selfishness "compatible" with the desired overall social outcome.

[5] Unfortunately, these detailed rules are sufficiently convoluted that it may be only the experimenter's word that motivates many subjects toward honesty, rather than the direct understanding that honesty is truly the best strategy.

[6] Two points should be remembered here. First, the term *revealed preference*, as it is used in economics, refers to a vast body of work produced over the past 70 years, complete with a collection of axioms relating choice patterns to rationality. Here we do not mean to invoke that entire literature, but only Samuelson's original observation that preference could be inferred from choice. Second, in some sense, even the incentive compatible self-report mechanism mentioned above is a form of revealed preference, insofar as casting the task as an auction causes subjects to "reveal" their valuations through their bids rather than directly "reporting" them to the experimenter. Moreover, the model-based methods we consider in the following section leverage revealed preference in that behavior is the medium from which valuations are reconstructed. In this section, we restrict ourselves to what might be called *naked* revealed preference – simple efforts to estimate valuation by directly tabulating choices across pairs of options.

[7] We avoid the use of more common terms: V for value, U for utility, EV for expected value, or EU for expected utility. Each of these terms has a very specific meaning, and here we aim to be as general as possible.

[8] Assumptions about how values combine must be validated by comparing predicted choice to observed choice, as we detail later in "Validating behavioral performance."

the neural representations of subjective desirability. A chief advantage of this approach over the methods discussed so far is that it is readily applicable to the study of value-based decision making in non-human animals. For example, it has been successfully employed in insolating neural decision variables in monkeys choosing between different quantities and flavors of fruit juice (Padoa-Schioppa and Assad, 2006).

An unfortunate limitation of many such simple revealed-preference designs is their assumption of *stationarity*. Because these methods generally rely on combining data across many trials to compute indifference points, they typically assume that the decision variables summarizing an option are identical every time that option is presented[9]. This stationarity assumption is clearly violated in at least two situations. The first is the case of satiation. Suppose, for example, that a subject chooses repeatedly between n pieces of chocolate and m pieces of apple. The marginal value of an additional piece of chocolate may well depreciate after having chosen chocolate several times in a row. Such stimulus-specific satiety effects (still being hungry overall, but being tired of chocolate in particular) can distort indifference points, making the inference of decision variables less reliable[10]. In humans, this difficulty can be alleviated by delivering a reward for one and only one of the subject's choices[11]. Subjects are typically warned that they will receive no rewards during the experiment, but that one trial will be randomly selected at the end of the experiment and their decision on that trial honored. Unfortunately, this trick is unavailable to anyone working with non-human animals, and thus is of little use to neurophysiologists.

Dynamic environments present a second setting in which the assumption of stationarity implicit in simple revealed preference becomes problematic. Increasingly, researchers have sought to enrich their experiments by varying contingencies between actions and outcomes over time, typically without warning subjects when and how the contingencies change. This technique is common in reversal-learning tasks, exploration–exploitation tasks, some foraging tasks, and any game-theoretic tasks in which an opponent reacts to the subject's actions – see, for example, Paton *et al.* (2006), Daw *et al.* (2006), Sugrue *et al.* (2004), and Barraclough *et al.* (2004), respectively. Often the experimenter is most interested in tracking changes in a decision variable during transition periods when a subject is actively adapting to the changing environment. Because they depend on combining information over many trials to find indifference points, simple revealed-preference methods can only provide precise estimates of decision variables once the subject has reached equilibrium with the environment, and not during these interesting dynamic periods.

The limitations of incentive compatible self-report and simple revealed-preference methods render them unsatisfactory for many experiments, particularly studies that involve either dynamic environments or non-human subjects. As a result, it is becoming increasingly common for researchers to construct explicit mathematical models of the decision process to estimate decision variables in experiments not amenable to the direct methods we have already discussed.

MODEL-BASED APPROACHES

The alternative to direct methods for estimating decision variables is the model-based approach,[12] in which a concrete mathematical model of the decision-making process is constructed (reviewed in Corrado and Doya, 2007; O'Doherty *et al.*, 2007). The model includes explicit decision variables that can be calculated on every trial; these quantities can then be used as *proxies* for the state of the subject's true internal decision variables. If these proxy variables co-vary with the subject's true decision variables, we should be able to localize the neural correlates of the subject's decision variables using the traditional correlative techniques of systems neuroscience[13].

---

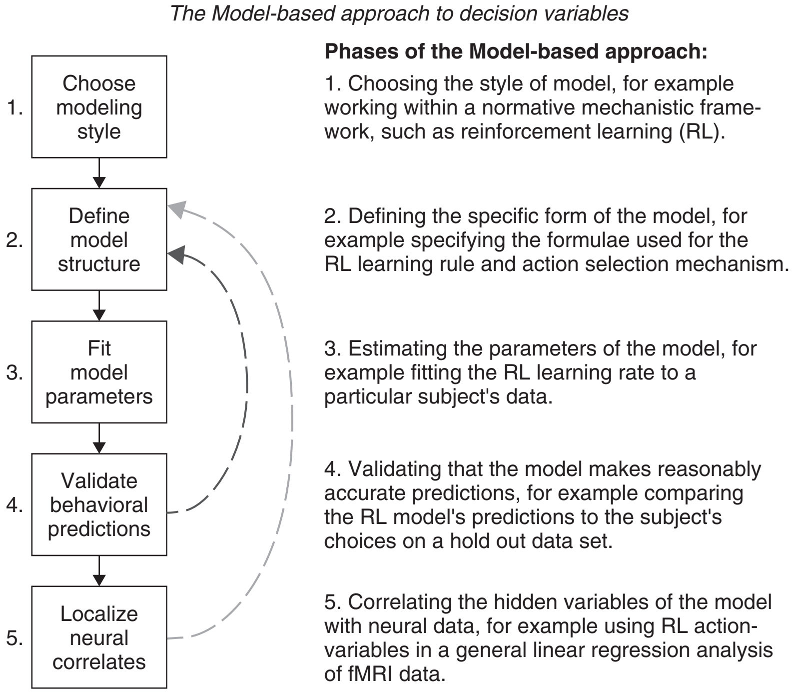
[9] It is in principle possible to include the effects of non-stationarities when deducing revealed preferences, but only with an explicit model of how preferences evolve in time, which falls under the rubric of model-based approaches considered more fully in the next section.

[10] Some clever researchers have actually turned stimulus-specific satiety effects to their advantage, counting on being over-fed on a particular food item to devalue it (e.g., Kringelbach *et al.*, 2003) – but more often than not satiety, or more generally, interactions between outcomes on different trials, is an annoyance.

[11] This trick is commonly employed in experimental economics to (1) avoid having to pay subjects large amounts of money for performing many trials and (2) avoid the risk that subjects will attempt to assemble a portfolio or advantageous combination of options over many trials.

[12] Sometimes called "Model-based fMRI" when fMRI is the experimental measure.

[13] Some researchers view the models they construct for this purpose as literal hypotheses about the details of neural computation. We prefer to view them as a means to an end: providing proxy variables of use in identifying decision circuits in the brain even if the details of computation differ substantially between the model and the neural implementation. This conservative stance frees us from the necessity of demonstrating that all elements of the model are plausibly implemented in the brain, and instead allows us to focus on our primary objective – identifying neural correlates of the key decision variables.



Phases of the Model-based approach:

1. Choosing the style of model, for example working within a normative mechanistic framework, such as reinforcement learning (RL).
2. Defining the specific form of the model, for example specifying the formulae used for the RL learning rule and action selection mechanism.
3. Estimating the parameters of the model, for example fitting the RL learning rate to a particular subject's data.
4. Validating that the model makes reasonably accurate predictions, for example comparing the RL model's predictions to the subject's choices on a hold out data set.
5. Correlating the hidden variables of the model with neural data, for example using RL action-variables in a general linear regression analysis of fMRI data.

FIGURE 30.3 The five phases of a model-based approach for investigating neural decision variables. Research programs tend to progress through these stages in the presented order. Importantly, the results of both the behavioral validation (4) and the neural correlations (5) steps can lead to revision of the model to better capture the data (dashed lines).

The general strategy for model-based investigations proceeds in several stages (Figure 30.3), each of which we will consider in a separate section below. For the remainder of this section we will focus on the first three steps in this process: selecting a general modeling framework and style, tailoring a specific model to the experimental design and scientific question, and fitting the model parameters to behavioral data. The fourth and fifth stages of the process – validating behavioral predictions and correlating model proxy variables with neural data – are common to both model-based and direct methods, and are examined in the subsequent sections.

Styles of Modeling

There are a great many styles of models and modeling frameworks. Rather than enumerating all of the possibilities, we will outline the most important dimensions in which modeling frameworks vary and give a few examples of frameworks that have been used successfully to model behavioral choice.

Aggregate vs Mechanistic Models

The principle axis of variation concerns the resolution at which a model describes behavior. Some of the most influential decision models from psychology and economics deal only with average behavior – for example, Herrnstein's Matching Law (Herrnstein, 1961) or Nash Equilibria (Nash, 1950). Although these models may be very successful in describing the *aggregate* properties of choice in a wide variety of circumstances, they do not say much about individual choices. To be useful for neurobiological purposes, models need to be *mechanistic*, meaning that they should explicitly detail how internal variables are updated and used to render choices on each trial. The brain must formulate and execute each choice individually, and the value of each decision variable must therefore be specified for each trial[14]. Fortunately, aggregate models can often be made mechanistic with a few simple extensions. For example, Herrnsteins's Matching Law can be implemented as a local mechanistic process in which animals allocate choices based on the ratio of their recent reward experience (e.g. Herrnstein and Prelec, 1991; Sugrue *et al.*, 2004). As we will see later in the chapter,

---

[14] Aggregate behavior is interesting, but the relevant neural mechanisms for understanding aggregate behavior are exactly those that mediate individual choices. Aggregate behavior is after all nothing more than the cumulative *result* of the decision maker's individual choices made one at a time.

a key feature of mechanistic models is that they often explicitly dissociate the option evaluation stage and the action selection stage (Figure 30.5).

Descriptive vs Normative Models

A second important axis of variation among decision models is *descriptive* vs *normative*. Descriptive models seek only to *describe* the behavior as it appears in the data (e.g. Herrnstein's Matching Law), whereas normative models *prescribe* a pattern of behavior for the particular task at hand. Under the assumption that the subject is a rational agent, normative prescriptions generally aim to realize some form of optimality, such as maximizing expected reward (as in Nash Equilibria). Of course, such notions of optimality depend critically on the assumptions used to derive them, which may not align well with the constraints imposed by natural environments. For example, correctly solving for Nash Equilibria requires that a player has complete information about the precise consequences of every available action and how these outcomes depend on the actions of other agents[15]. Unfortunately, in many real-world situations sufficiently accurate information about what to expect from the world or one's adversaries is unavailable, making the standards set by Nash unattainable for decision-making systems operating with limited information, time, and computational resources. This fragile dependence of "optimality" on specific assumptions is one of several reasons why researchers do not generally insist on limiting themselves to normative frameworks when constructing models of behavioral choice. In most cases the model that provides the more accurate account of the data is preferable, even if that model implies "rough-and-ready" decision making on the part of the subject.

These two dimensions – aggregate/mechanistic and normative/descriptive – make up the principle axes of variation in models of decision making. Figure 30.4 shows a table of example models in each of the four combinations of subtypes. Many other examples exist, and within each category there are additional important axes of variation. In the next two sections we will develop an example mechanistic-descriptive model of a decision mechanism. In doing so we will describe two additional axes of variation – the number of parameters a model requires, and whether the model makes probabilistic or deterministic predictions.

Specifying the Details of the Model

After selecting a style of model to use, we must specify its detailed mathematical form. This is largely a creative process, one that lacks a rote procedure for arriving at the best model for a particular experiment. We will consider a simple but plausible example study, and explore a few possible decision models that could be used to compute proxy decision variables and predict behavior in this hypothetical study.

In our hypothetical example experiment, we will require subjects to choose between two "lotteries." In the simplest case, a lottery is an offer of receiving reward x with probability p – for example, a 10% chance of winning \$30.[16] For our example model, we will elect to work within the mechanistic and descriptive quadrant of model frameworks (Figure 30.4, upper right quadrant). As illustrated in Figure 30.5, specifying a mechanistic model requires us to define both an option evaluation mechanism and an action selection mechanism. The mechanism selected for option evaluation is obviously critical, because this stage is directly responsible for computing the proxy decision variables that we will ultimately seek to correlate with neural activity. However, we will see below that the choice of action selection mechanism can also have substantial (if indirect) effects on the imputed decision variables that we calculate.

In the following subsection we will consider several possible option evaluation mechanisms, and in doing so explore the pros and cons of varying the number of parameters incorporated into the model. We will then go on to compare deterministic and probabilistic action selection mechanisms to marry with our preferred option evaluation mechanism in order to construct a complete model of decision making for our lottery experiment.

Option Evaluation-the Pros and Cons of Parameters

The role of an option evaluation mechanism is to convert the features of options into decision variables that feed the action selection mechanism. For our example experiment, this amounts to defining the

---

[15] Even in cases where games that lack complete information are recast as Bayesian games, rich probability distributions over possible consequences and opponents' preferences are required. The correctness of the prescribed actions is sensitive to the accuracy of these large probability tables, which may themselves be intractably difficult for a decider to estimate and manipulate.

[16] Lotteries can, of course, be much more complicated – for example, a 10% chance of winning \$30, a 5% chance of winning \$100, a 1% chance of losing \$200, and an 84% chance of winning nothing – but here we restrict ourselves to the simplest case of winning x with probability p, and nothing with probability $(1 - p)$.

Example modeling frameworks

	Aggregate	*Mechanistic*
Descriptive	**Herrnstein's Matching Law** describes the average distribution of choices by animals foraging on multiple resource patches, without regard to whether that is the most efficient allocation of choices. (Herrnstein, 1961)	**Linear Nonlinear Probabilistic** models are three-stage feed-forward models that use a linear transformation to construct decision variables and a simple nonlinearity to make probabilistic predictions of choice. (e.g. Corrado et al. 2005)
Normative	**Nash Equilibria** are stable points of competitive games, prescribed sets of strategies for which no player can improve his average payoff by unilaterally changing his behavior. (Nash, 1950)	**Reinforcement Learning** models are simple mechanisms that update the values of states or actions based on past experience, converging to optimal values over time. (Sutton & Barto, 1988)

FIGURE 30.4 Examples of models in four corners of the descriptive/normative and aggregate/mechanistic space. Many other examples exist in each category, and the boundaries between categories may not be as crisp as implied. For example, as mentioned in the main text, Herrnstein's matching law can be made more mechanistic relatively easily. Similarly, some forms of reinforcement learning (e.g., Q-learning) can be shown to be optimal in restricted environments (stationary environments that allow the agent infinite time to explore), but their optimality cannot be guaranteed under more realistic conditions.

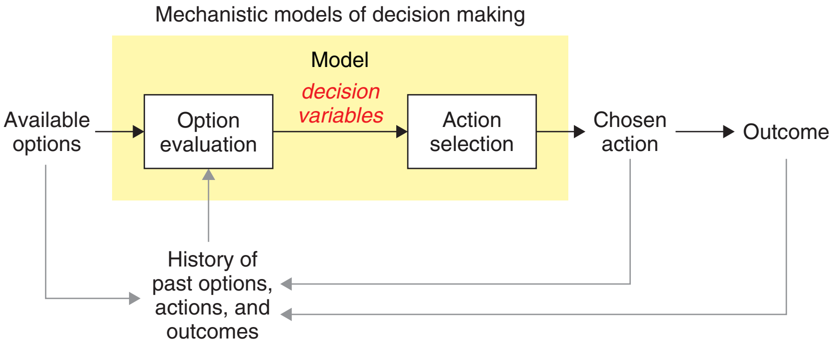


FIGURE 30.5 Mechanistic models of decision making. Mechanistic models of decision-making agents often have two separable parts: an option-evaluation stage and an action-selection stage. The option-evaluation mechanism responds to immediately available options, but may also incorporate information from the past (e.g., whether an option previously delivered a positive outcome). The action-selection mechanism weighs the desirability of the available options and renders the final decision. Whether that chosen action ultimately results in a positive or negative outcome depends on factors external to the agent.

subjective desirability of lotteries A and B in terms of their probabilities of payoff p, and the dollar amounts of their payoffs x:

$$(D_A, D_B) = g(p_A, x_A, p_B, x_B) \qquad (30.2)$$

It seems intractable to consider all possible functions g; we must therefore make some simplifying assumptions. Choosing which assumptions to impose is part of what makes defining a behavioral model both challenging and scientifically interesting. One common, if relatively strong, assumption is that options A and B are evaluated independently and by identical mechanisms. This greatly reduces the space of functions we are considering:

$$D_i = f(p_i, x_i) \quad i \in \{A, B\} \qquad (30.3)$$

In choosing the function f, the most important axis of variation is the number of parameters we use to describe the function, and thus our model. At one end of the spectrum, *parameter-free* models assert that f (and thus the behavior of the subject) can be described by a mathematical function of our choosing, that does not need to be tuned to fit the data at hand. Relaxing this assumption a little, *parametric models* assume a very specific functional form for f, but include a small number of free parameters that can be adjusted to fit the behavioral data of individual subjects. At the far end of the spectrum, *reconstructed models* take this flexibility one step further by making relatively modest assumptions about the form of f and attempting to estimate f directly from the data – in essence, by fitting a very large number of parameters to the data. To appreciate some of the advantages and disadvantages of these different model types, we consider some concrete possibilities in each category[17].

A parameter-free model offers the simplest starting point. Classical economic theory tells us that the worth of a lottery is just the expected monetary value of its outcome, and so we might hypothesize that the desirability of option A is proportional to the expected value:

$$D_A \propto EV_A = p_A x_A \tag{30.4}$$

where p_A is the probability that the lottery will pay amount x_A. In the parameter-free approach, no further specification of the model is necessary. We can compute px for any option we offer, and search for neural activity that covaries with these putative values. We do not even need to determine the constant of proportionality between D_A and EV_A because its value has no effect on statistical measures of correlation. Unfortunately though, we can rarely write down parameter-free accounts of behavior that fit real choice data with any precision. Indeed, given the many factors at play in biological systems, it is rare that behavior can be reduced to the type of an elegant mathematical formula you might find a on a T-shirt. This is not physics, after all.

In general, incorporation of even a few additional degrees of freedom improves a model's ability to describe behavior considerably. Such, *parametric models* express a strong hypothesis about the form of the decision variable, but allow individual variation in how decision variables are computed – variation that is characterized and absorbed by a small number of descriptive terms. For example, if we found that our parameter-free model (equation 30.4 above) did not fit the behavioral data very well, a neo-classical economist might suggest that desirability should be proportional not to the raw expected value but instead to the expected *utility* of an option:

$$D_A \propto EU_A = p_A U(x_A) \tag{30.5}$$

where U is the individual subject's utility function. The utility function can capture the fact that equal increments in the size of a reward may not be equally desirable to the subject – for example, some subjects may work hard to win \$1 as opposed to zero, but exert little effort to win \$101 as opposed to \$100. If we assume that a reasonable utility function should be monotonically increasing and free of inflection points, we might consider using a power law to characterize its shape:

$$U(x) \approx x^{\alpha} \tag{30.6}$$

where values of α less than one describe concave functions and values of α greater than one describe convex functions. This leads to a simple parametric model of desirability:

$$D_A \propto p_A x_A^{\alpha} \tag{30.7}$$

We can fit this parametric form directly to behavior; fine-tuning the value of α for different subjects may allow for a more accurate description of each subject's behavior[18]. We will address methods for fitting parameters to data later in the chapter.

Reconstruction techniques take this idea a step further. Instead of assuming a particular functional form for U and fitting the parameter α to the data, we might try to measure the shape of each subject's utility function directly from the data, and then use this estimated utility function $\hat{U}$ in our model:

$$D_A \propto p_A \hat{U}(x_A) \tag{30.8}$$

This approach is considerably more ambitious than the single parameter model described by equation (30.7). In essence, the "reconstruction" approach involves estimating a large number of parameters (for example, by expanding the function $\hat{U}$ as a polynomial), and therefore requires much larger data sets and clever experimental designs to be tractable (see Corrado *et al.*, 2005, and Lau and Glimcher, 2005, for successful examples of the reconstruction of subjective functions, though not utility functions over lotteries, as would be needed for this example experiment)[19].

Figure 30.6 shows a graphical representation of the option evaluation mechanisms we have considered. Moving from left to right, the number of assumptions decreases, while the number of parameters and flexibility

---

[17] Rather than three discrete categories, this is best thought of as a continuum, ranging from zero parameters at one extreme to a very large number of parameters (perhaps hundreds) at the other.

[18] Our discussion of how to model lottery valuation with a utility function is simplified for didactic purposes; for a serious treatment of the problem see, for example, Holt and Laury, 2002.

[19] Notice that equation (30.8) does not endeavor to reconstruct the two-dimensional function f in equation (30.3) in its entirety. In general, reconstruction models still employ simplifying assumptions to reduce the problem to a manageable size. In our case, we assumed the option evaluation mechanism to be linear in its dependence on p – a relatively strong claim common to all three of the models we have considered. While this assumption is common, there is nothing sacred about it. It might be that assuming a linear dependence on x would be a better model of behavior, i.e. $D = H(p)x$, rather than $D = pU(x)$. With sufficiently large data sets we might even get away with assuming only that the dependence on p and x is separable, i.e. $D = H(p)U(x)$. Any of these would be far easier than reconstructing the full function $D = f(p, x)$.

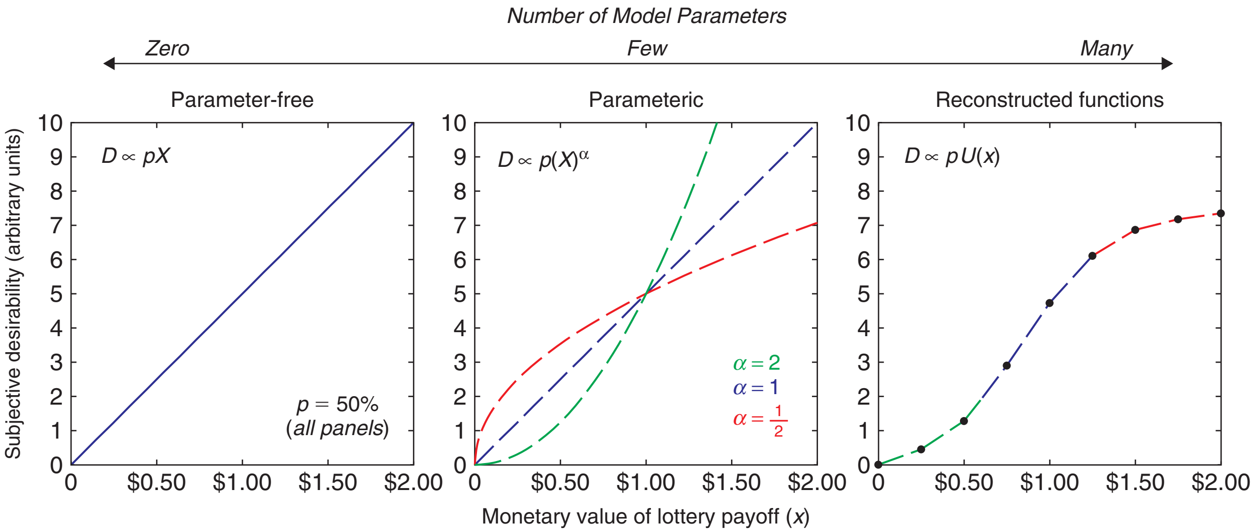


FIGURE 30.6 Example option evaluation mechanisms for the subjective desirability of a lottery option with varying numbers of model parameters. The model at the left has no free parameters, and hypothesizes that the subjective desirability of a lottery is directly proportional to its expected monetary Payout. The model at the center uses a single free parameter to characterize the value of a lottery as either an accelerating (green) or decelerating (red) function of possible payoff. The reconstructed function at the right uses a large number of free parameters to characterize the specific shape of the relationship between payoff and subjective desirability, allowing for richer descriptions of behavior at the cost of far greater data demands.

of the model increases. Unfortunately, with increasing flexibility comes a steadily rising demand for data to fit the increasing number of parameters. Not surprisingly, most successful research programs take the middle path – the parametric approach – avoiding both the extreme rigidity of the parameter-free models and the extreme data demands of function reconstruction.

The three example models comprised of equations (30.4), (30.7), and (30.8) provide explicit formulae for constructing decision variables that capture the desirability of a lottery option. Each example model is a reasonable candidate for the option evaluation process diagrammed on the left side of Figure 30.5, and satisfies our principal goal of explicitly specifying the putative decision variable. However, to fit the parameters of the model to data and to validate the model's behavioral predictions, we must be able to compare the subject's actual choices with choices predicted by our model. To make this comparison, we must complete our model by adding a concrete action selection mechanism, as diagrammed on the right-hand side of Figure 30.5.

Action Selection: Deterministic or Probabilistic Predictions

The role of the action selection stage is to convert putative decision variables into concrete predictions about choice. Because the action selection stage is downstream of the decision variables, it has less effect on our imputed decision variables than does the option evaluation stage. It exerts some influence, however, because it impacts the *parameters* of the option evaluation stage that fit the data best – thus affecting the computation of putative decision variables via changes in those parameters. Moreover, it has a critical role in our overall assessment of how well the model predicts choice, and thus our level of confidence in the hypothesized decision variables.

Action selection mechanisms come in two basic flavors: those that make *deterministic predictions* and those that make *probabilistic predictions*. By far the most common deterministic mechanism is to predict that the subject will choose the option with the greatest subjective desirability[20]

$$\textit{chosen} = \arg\max_{i}(D_i) \qquad (30.9)$$

This action selection mechanism is extremely popular in economics, offering the advantages of simplicity and clear interpretability.

In truth, however, subjects rarely choose the "best" alternative in a deterministic manner. For this reason, many models incorporate a probabilistic selection criterion: predicting simply that the *probability* of choosing an option increases with its desirability. Sigmoidal functions are generally used to implement probabilistic

---

[20] For those unfamiliar with the *argmax* notation, this equation states that *chosen* is set to the argument i that maximizes the value of the expression D_i, e.g. amoung two options, A and B $\textit{chosen} = A$ if $D_A > D_B$.

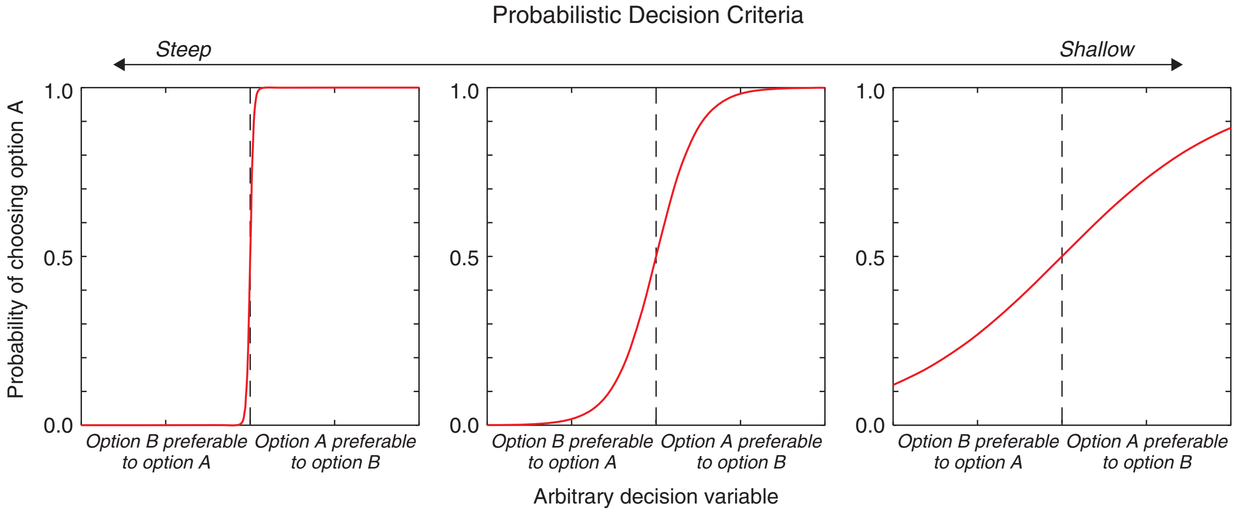


FIGURE 30.7 Probabilistic decision criteria. Probabilistic decision criteria vary in how steeply they increase as a function of the desirability of an option. For example, the steep criterion at the left predicts that the probability of choosing option A is near 1 when option A is only slightly more desirable than option B. This sharp discrimination between the relative desirability of options A and B approximates the behavior of a simple deterministic function that always selects the most desirable option. The sigmoidal function at the center shows a smoother transition between exclusively choosing A or B. The gradual function at the right biases choices in favor of the more desirable option, but never chooses either option exclusively.

selection criteria, and one of the most commonly used is the *softmax* function:

$$pchose(i) = \frac{\exp\{\kappa D_i\}}{\sum_j \exp\{\kappa D_j\}} \tag{30.10}$$

where κ describes the steepness of the relationship between desirability and probability. Figure 30.7 shows the shape of the softmax function for different values of κ. The left-hand panel depicts a function that is so steep as to effectively choose the more desirable option 99.9% of the time. In contrast, the right-hand panel illustrates a function that introduces only modest biases in favor of the more desirable option, even for large differences in option desirability.

Probabilistic action selection stages allow for the possibility of noise or variability in choice behavior, and therefore encompass a far greater range of possible decision mechanisms – but, critically, they do so at relatively little cost. At the limit, probabilistic models with very steep decision criteria subsume deterministic models, accommodating behavior even when the data are best described by a nearly perfect mapping from desirability to choice.

To define a complete choice model, we can combine either of these example decision functions, equations (30.9) or (30.10), with any of our three example evaluation functions, equations (30.4), (30.7), or (30.8). For example, using equations (30.7) and (30.10), we can write down one possible model as:

$$pchose(i) = \frac{\exp\{\kappa p_i x_i^{\alpha}\}}{\sum_j \exp\{\kappa p_j x_j^{\alpha}\}} \tag{30.11}$$

This equation describes the probability that the subject will choose lottery i, which has probability p_i of paying x_i, given the parameter α describing the concavity of the subject's utility function and the parameter κ describing the steepness of the subject's decision function.

Parameter Estimation

Having seen that most useful models will involve some number of parameters, we now consider briefly how best to estimate those parameters from the data. In general, this is done by optimizing some *objective function*, a measure of the goodness of fit between the model and the data. Essentially, this means comparing the actual decisions made by the subject to the decisions predicted by the model, and finding the set of parameters that minimizes the discrepancy between the two.

Having specified a model that makes concrete predictions about choice – for example, equation (30.11) – we can find the model parameters that best describe a subject's behavior by optimizing our objective function. The two most common techniques are to maximize the likelihood of the data given the model's predictions

$$L = \prod_t pchose_t(x_t) \tag{30.12}$$

or to minimize the squared error between the data and the model's predictions

$$E = \sum_t \left(1 - pchose_t(x_t)\right)^2 \tag{30.13}$$

where, in both cases, $pchose_t(x)$ is the model's predicted probability that the subject will choose option x on trial t, and x_t is the option actually chosen by the subject on trial t. Maximum likelihood (ML) and minimum squared error (MSE) parameter estimates are often very similar, and rarely is there a strong case *against* using either method. MSE parameter estimates are often easy to compute, and can accommodate both deterministic and probabilistic models. On the other hand, many view ML as the most natural sense of "goodness of fit" for probabilistic models. Moreover, in cases where there is a strong *a priori* reason to give preference to some choices of model parameters over others, ML can be extended into more advanced *Bayesian* methods (e.g. Samejima *et al.*, 2004). The details of the procedures for actually arriving at ML, MSE, or Bayesian parameter estimates depend on the details of the model being employed and are beyond the scope of this chapter, but a great many references and software tools are available (e.g. van der Heijden *et al.*, 2004).

VALIDATING BEHAVIORAL PERFORMANCE

Whether we arrived at our putative decision variables through one of the direct or model-based approach, it is critical to assess the performance of those decision variables in predicting behavior. We are unlikely to achieve our ultimate goal – identifying neural correlates of hidden decision variables in the brain – if our postulated decision variables do not describe observed behavior accurately.

For those who have taken a model-based approach, assessing behavioral performance is easy: compare the choices (or probability of choice) output by the model with the subjects' actual choices and look for discrepancies. Putative decision variables arrived at through direct methods require some additional machinery to make their predictions testable. In essence, they require an action selection mechanism just like those outlined above to transform their raw decision variables into concrete predictions about actions. Often, for the sake of simplicity, this is achieved using the deterministic "choose-the-best" rule as an implicit action selection stage (equation (30.9)), but there is no reason why a probabilistic mechanism could not be used.

Simple Predictive Performance

The most obvious metric of a model's performance is its ability to predict a subject's choices. With a few modifications, the objective functions that we employed to fit models earlier in the chapter (see "Parameter estimation") can also be used to quantify predictive performance. Clearly, the same data should not be used to define the decision variables *and* to assess the accuracy of their predictions. Asking a model to predict data that were used to set its parameters conveys an unfair advantage, and would lead to inflated confidence in the model's performance. Thus, separate subsets of the data must *always* be used to fit parameters and assess predictive performance[21].

For action selection mechanisms that make probabilistic predictions, the likelihood function given in equation (30.12) is by far the most commonly used metric of performance. For data sets of any significant size, the raw likelihood will be an absurdly small numerical value, because it is obtained by multiplying a long list of numbers, all less than one. For this reason, most researchers report (and indeed optimize) the log-likelihood, by taking the logarithm of both sides of the equation:

$$LogL = \sum_t \log pchose_t(x_t) \tag{30.14}$$

where, as before, $pchose_t(x)$ is the model's predicted probability that the subject will choose option x on trial t, and x_t is the option actually chosen by the subject on trial t. For the model-based approach, when comparing the performance of models having different numbers of parameters, heuristics like Bayesian Information Criterion (BIC) or Akaike's Information Criterion (AIC) can be used to augment equation (30.14) with a term that penalizes models for having

---

[21] The roles of the subsets may be exchanged, or even rotated in more complex schemes, and the results combined to produce so-called *cross validated* estimates of predictive performance; see Hastie *et al.*, 2001.

additional parameters[22] (see Burnham and Anderson, 2002, for a tutorial).

Likelihood metrics like equation (30.14) cannot be applied to action selection mechanisms that make deterministic predictions. Fortunately, error metrics like equation (30.13) are available as ready substitutes. For deterministic predictions, equation (30.13) reduces to the percentage of the subject's actual choices that are correctly predicted by the model, a simple but very intuitive metric (e.g. "our model correctly predicted 98% percent of subjects' decisions"). While commonly used for deterministic predictions, this metric is rarely used for probabilistic models because it fails to account for the confidence with which predictions are made.

Establishing Upper and Lower Bounds on Predictive Performance

The metrics of predictive accuracy described above can quantify how well a particular model[23] describes a specific data set, and thus can be used to compare the performance of competing models. However, thus far we lack a "gold standard" that specifies in principle how well a model could describe the data. More generally, we would like to establish bounds on both the low and the high ends of performance, so as to assess any specific model relative to random and ideal performance.

We can establish a reasonable lower bound on performance by applying a metric of predictive performance to a set of "random" predictions. If a model cannot out-perform a random noise generator, clearly we should discard it and search for something more promising[24]. Establishing a reasonable upper bound on predictive performance is more difficult. Predicting every choice perfectly is, of course, an absolute upper bound, but an unreasonable expectation for any sizeable data set. More sensibly, for the data under consideration, we would like to be able to compare the predictive performance of the model being considered with the maximum achievable performance of any compact model.

An elegant approach to calculating maximum achievable performance was suggested to us by William Bialek of Princeton University. For any model M using the inputs X to predict choices C, we can establish an upper bound on the quality of the predictions using Shannon's mutual information:

$$I(C;M(X)) \leq I(C;X) \tag{30.15}$$

Put into words, the information about subject choices contained in the model's predictions cannot exceed the information about those choices contained in the model's inputs (Figure 30.8). Models cannot add Shannon information; they can only distill it. This realization offers a practical method for calculating an upper bound on predictive performance. Given particular inputs (for example, descriptions of the options A and B in terms of probability and payoff) we can quantify how well *any* model relying on these inputs could possibly do by computing $I(C;X)$. If a subject behaves consistently, making similar patterns of choices for repeated presentations of identical option sets, the upper bound will be high, perhaps approaching (if never reaching) the elusive goal of perfect prediction. If, on the other hand, the subject behaves erratically, making different choices for repeated

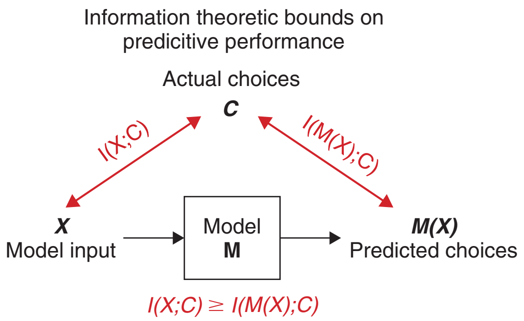


FIGURE 30.8 Information theoretic bounds on predictive performance. Information theory provides a rigorous bound for the predictive performance of behavioral models. Consider all models M that use the variables X as inputs. The mutual information between the model's predictions, $M(X)$, and the actual choice data, C, cannot exceed the mutual information between X and C. Thus, $I(X;C)$ provides an upper bound on performance for all models that use X to predict C.

---

[22] These metrics include a penalty for additional parameters because we would expect any model with more parameters to fit better. Unfortunately, BIC and AIC are only rough heuristics for making this correction.

[23] The same metrics can be used to evaluate decision variables arrived at through direct (rather than model-based) approaches; however, as mentioned earlier, the directly inferred decision variable must be coupled with a particular action selection mechanism before its predictive accuracy can be assessed.

[24] There is some liberty in how we define these "random" predictions. The simplest option is to make *choices* that are random: a dummy set of predictions, *pdummy*, uniform and independent of the options offered: $pdummy_t(k) = 1/m$ for all trials t and options k when choosing between m alternatives. Another possibility is to make *random predictions*, i.e. drawn at random from a uniform distribution: $pdummy_t(k) = r$ and [0,1]. For the purpose of comparison with a single particular model, there is also the possibility of using *randomly shuffled predictions*, that is to use predictions that are drawn at random from the predictions made by the model under scrutiny: $pdummy_t(k) = pchose_j(k)$, so that the t-th dummy prediction is set to be the comparison model's j-th prediction for a randomly chosen trial j.

presentations of identical option sets, the upper bound on model performance will be lower. For example, in our own dynamic foraging experiments, the history of a subject's recent rewards and choices contained nearly 0.50 bits of predictive information about their next choice. This set an upper bound on the performance of any model that used reward and choice history as an input to predict a subject's behavior. The mechanistic models considered in Corrado (2007) yielded decision variables that retained between 50% (the worst models tested) and 90% (the best models tested) of this available information.

This procedure for establishing upper bounds on performance has a serious limitation: the size of the data set required to estimate $I(C; X)$ is typically much larger than can be gathered from individual human subjects in a reasonable amount of time. The reason is that the probability table required to map out the function $I(C; X)$ over all possible inputs X and range of choices C is usually very large. Our data set, for example, included tens of thousands of trials per subject, and still did not permit us to specify $I(C; X)$ completely. Thus, though the idea shows great promise, the problem of limited data sets currently diminishes its practical generality.

Generative Performance

In contexts where successive decisions are made independently of one another (i.e. the subject would have made the same decision on a particular trial even if the trials had been presented in some other order), evaluating average predictive performance is frequently sufficient to quantify behavioral performance. However, in contexts where the environment (or the state of the subject) has interesting temporal dynamics[25], an additional assessment is needed: the ability of a model to *generate* realistic patterns of behavior in a simulated experiment.

To highlight the distinction between *predictive* and *generative* performance, consider a television weatherman who everyday forecasts, "today's weather will be the same as yesterday's." Because of temporal correlations in weather patterns, this algorithm actually *predicts* tomorrow's weather correctly on an impressively high proportion of days. However, because the algorithm incorporates no understanding of the mechanisms underlying weather production, it will fail miserably at *generating* realistic patterns of weather change over time. Thus, adequate predictive performance (as described in the previous section) is no assurance of generative success.

Metrics for evaluating generative performance are more difficult to define than their counterparts for assessing predictive performance. The principle, however, is simple: compare real data with synthetic data generated using the model, and look for discrepancies. In practice, of course, we should focus on dynamics pertinent to the particular task. For example, in a reversal-learning task with an abrupt unsignaled change in reward contingencies, we might look at the time-course of the subject's adaptation to the new conditions (e.g., Kennerley *et al.*, 2006). On the other hand, in a foraging task we might compare the distribution of the lengths of time spent at each foraging patch (Corrado *et al.*, 2005). Regardless of the specifics, the ideal model will produce synthetic data that are statistically indistinguishable from real data. We are not aware of any simple formula that specifies the precise meaning of "statistically indistinguishable" for the completely general case. Instead, evaluating generative performance is more akin to a "Turing test"[26] for behavioral models: if experimenters using any of the behavioral measures of interest cannot tell which data are generated by the model and which are generated by the subject, then the model is presumably a good one.

THE QUEST FOR NEURAL CORRELATES

The traditional program of experimental psychology and economics ends here, with the extraction of internal decision variables that have demonstrable predictive power. Neuroscience and neuroeconomics aim to go further: to use these putative decision variables to elucidate decision-making circuitry within the brain. The core idea is to use these inferred variables as proxies for the subject's true decision variables in a correlative study (Figure 30.2b). We have already

---

[25] Dynamic environments include any task where contingencies between actions and outcomes change over time and must be inferred by the subject, as is common in foraging tasks, reversal-learning tasks, exploration–exploitation tasks, and nearly all game-theoretic tasks. Dynamics internal to the subject refer to any setting in which the subject's next choice depends either on previous actions or previous outcomes, as in contexts involving reinforcement learning or "reactive" player models.

[26] Alan Turing made a famous proposal that a reasonable test for artificial intelligence would be to stage what amounts to an instant messaging text conversation; if humans were unable to distinguish between talking to another human or to a machine, then that machine could be said to be intelligent (Turing, 1950).

alluded to studies in which decision variables recovered through the direct methods described earlier in the chapter have been used to probe underlying neural signals (see, for example, Plassman *et al.*, 2007, for incentive compatible direct report in human fMRI; Padoa-Schioppa and Assad 2006 for revealed preference in monkey electrophysiology). In this section we consider three examples of the integrated model-based approach and its successful application to the analysis of neural signals (Figure 30.3).

The three example studies differ in the physiological signals measured, the species studied, and the specific form of the model used to capture the behavioral data. However, all employ tasks in which option values are uncertain and change unpredictably through time. In each study the authors built a mechanistic model of task behavior that had distinct option evaluation and action selection stages (Figure 30.5), and used the model's internal decision variables to identify neural correlates. Finally, and perhaps most importantly, in each case the authors ultimately refined their initial model(s) in response to subsequent behavioral or neural findings. This last point emphasizes the inherently iterative nature of this process (dashed lines in Figure 30.3), and the fact that a behavioral model, regardless of its normative elegance or predictive power, should not be regarded as anything more or less than a working hypothesis (but see Chapter 3 in this volume for an alternative view).

Matching Behavior and Value-based Action Selection in Area LIP

In one of the earlier examples of the model-based approach (Figure 30.3), the authors of this chapter employed a Linear-Nonlinear-Poisson (LNP) framework to build a parametric model of the behavior of rhesus monkeys engaged in a dynamic foraging task (Sugrue *et al.*, 2004). This model was inspired by, and amounts to, a mechanistic implementation of Herrnstein's Matching Law (Herrnstein, 1961). At its evaluation stage the model uses simple exponential filters to compute a weighted average of recent rewards from each option, while at its action selection stage it replicates the matching law by using divisive normalization to transform these value estimates into a probabilistic output appropriate for selecting an operant response.

Despite its simplicity, this model was successful both in predicting the animals' choices and in generating realistic simulated behavior. Furthermore, the model's trial-to-trial estimates of value revealed a monotonic encoding of option value in the firing rates of individual neurons in cortical area LIP, a region known to be involved in the high-level planning of eye movements. Interestingly, while option color was the cue to value in this task, the map of option value in LIP was in spatial coordinates, appropriate for the probabilistic selection of an appropriate action on the current trial. This transformation from color-based to spatial encoding of value suggested to the authors that activity in LIP reflects, and might even effect, the action selection mechanism that maps value to action in this task (Figure 30.5).

Corrado *et al.* (2005) later took advantage of the large datasets collected in this experiment to move from a parametric model to one that reconstructed the functions operating at the option evaluation and action selection stages directly from the data. (Figure 30.6) The resulting model differed in a number of respects from the original parametric model, the evaluation stage being better characterized by a double rather than a single exponential, and the action selection stage by a differential (softmax) rather than a divisive operation. Using the validation techniques discussed above, the authors showed that the reconstructed model provided a better account of both animal behavior and LIP firing rates than the original parametric model, in spite of the fact that it no longer bore any formal resemblance to the original matching law. This result highlights the potential for the choice of modeling style – in this case, parametric versus reconstructed – to influence the conclusions of the entire behavior–model–physiology process.

fMRI Correlates of Reinforcement Learning Models

The next two studies used Reinforcement Learning (RL) as the basis for the evaluative stage of their behavioral models (Sutton and Barto, 1988). Accordingly, in addition to variables that represented the learned values of competing options, these models included a RL prediction-error term that quantified the difference between the value of the chosen option and the outcome of that choice. The correspondence between the error signals postulated by RL models and the phasic responses of midbrain dopamine neurons recorded in simple conditioning tasks has led to the proposal that RL might reflect the literal mechanism through which the brain learns the values of states and actions (see, for example, Schultz *et al.*, 1997; also Part 3 of this volume). In that regard, the following studies are of particular interest because they are among the first neuroimaging studies to examine RL in choice situations where subjects are actively engaged in learning option values and using those values to guide behavior.

Daw, O'Doherty and colleagues (2006) employed a well-characterized behavioral paradigm called the "N-armed bandit" task. Their subjects made repeated choices among four slot machines, each paying out with an undisclosed probability that varied unpredictably through time. This design placed their human subjects in a situation similar to that of the monkeys in the matching task described above: challenged with estimating the machines' relative values, and using those estimates to decide which machine to choose on each trial. An RL algorithm provided the value estimates for the authors' behavioral model, while a softmax operation comprised the action selection stage (softmax having been selected after considering a variety of different rules using maximum likelihood techniques).

The authors used the internal variables of the resulting model as regressors to analyze fMRI BOLD measurements made while their subjects performed the task. The task-related fMRI signals identified in their analysis initially seemed to confirm what other groups had reported previously: BOLD signals in ventromedial prefrontal cortex (vmPFC) correlated with the model's estimates of the relative value of the chosen option, while signals in the striatum correlated with the error term of the RL algorithm (Knutson *et al.*, 2000; McClure *et al.*, 2003; O'Doherty *et al.*, 2003; see also Chapter 24 in this volume).

If it had ended there, this result would have marked an important demonstration of RL signals within the brain in the context of a task in which those signals have true behavioral relevance. However, the authors extended their analysis in a further and highly revealing direction. Using the predictions from their model's action selection stage, they classified each choice made by their subjects as either exploitative (choosing the slot machine assigned the highest value by the model) or exploratory (choosing a lesser valued option). They then searched for regions where the BOLD signal differentiated between these two classes of trials, identifying areas of anterior frontal and posterior parietal cortex that were preferentially active during exploratory decisions. This result suggested the involvement of distinct brain regions that help guide exploratory actions, above and beyond the signals expected in the most naïve forms of RL.

Recent work by Hampton and colleagues (2006) extended this result in an experiment that is among the first in the field to use neural signals to explicitly test competing behavioral models. These investigators studied behavior and fMRI BOLD responses in the context of a simple probabilistic reversal-learning task. Human subjects could improve their performance through their knowledge of the structure of the task, which included quasi-periodic reversals in the identity of the more valuable of two available options. Appropriately, the authors found that a state-based Bayesian Markov model that incorporated information about this higher-order task structure, provided a better account of subjects' behavior than simple RL models that were blind to such contextual information. Importantly, the decision variables at the evaluation stage of these two competing models sometimes diverged widely in their predictions, allowing the authors to show that BOLD signals in the striatum and vmPFC – two areas associated with RL signals – more closely tracked the update and value signals derived from the abstract states of their Bayesian model. To our knowledge, this is the first neuroeconomics study to use neural signals to distinguish between two candidate models of decision making. However, as our understanding of neural mechanisms matures and our models increase in sophistication, such hypothesis-driven approaches should become routine.

CONCLUSION

We began this chapter with the challenge of studying the neurobiology of decision making in settings where we lack easy access to the decision variables that guide choice. We subsequently saw how mechanistic behavioral models can offer a solution to this problem by providing reasonable proxy variables with which to correlate physiological signals. The putative decision signals identified through this approach must ultimately become the raw materials for a neurophysiological program aimed at transforming our behavioral models from preliminary sketches to plausible blueprints for the brain's actual decision-making machinery.

Such a program entails two phases. The first demands the detailed characterization of neural signals from the perspective of different hypotheses and models. As discussed previously, an active dialogue between modeling and physiology is essential for this process to converge on a representation that is true to the underlying neural signals. The second phase hinges on the critical distinction between correlation and causation. Demonstrating neural correlates of a decision variable is, in principle, straightforward; it is substantially more challenging to prove that the correlated neural activity plays a causal role in the brain's decision-making process in the manner suggested by the proposed decision variable.

Establishing causal links between neural responses and perceptual or cognitive phenomena is a fundamental challenge faced by researchers not only in neuroeconomics, but in all of cognitive neuroscience. Historically, support for links between anatomy and function has come from patients or experimental animals with lesions restricted to the anatomic area of interest. Indeed, lesion studies first implicated vmPFC in value-based decision making by demonstrating that damage to this region impaired performance on reward-cued reversal learning tasks (reviewed in Rolls, 2000) and other tasks in which the best choice on each trial had to be inferred from the outcomes of earlier choices (see, for example, Bechara *et al.*, 2000).

Although they have proven useful for certain purposes, lesion techniques are rather blunt tools for dissecting neural mechanism. What we want are techniques that are temporally, spatially, and functionally precise, allowing us to manipulate, selectively and reversibly, the activity of a restricted population of neurons while leaving the rest of the system intact. Toward this end, transcranial magnetic stimulation (TMS) has been used in humans to induce transient inactivations of particular brain areas, permitting inference about their involvement in a given perceptual or cognitive function (reviewed in Hallett, 2007). In animals, similar reversible inactivations can be achieved through the local application of drugs (such as muscimol or lidocaine) which briefly silence neural activity in a restricted region (e.g. Hikosaka and Wurtz, 1985; McPeek and Keller, 2004).

Causal relationships between neural activity and cognitive function can also be tested using intracortical microstimulation (ICMS) – a technique that has relatively high spatial, temporal, and functional resolution. The idea behind this technique is simple: characterize the response properties of a local population of neurons, place an animal in a situation in which the signals from that population have direct relevance to behavior, and then ask whether microstimulation of the neurons influences behavior in a predictable manner. Pioneered in the field of perceptual decision making (e.g. Salzman *et al.*, 1990), this technique has more recently been applied successfully to the study of visual attention, face discrimination, target selection, and reward-based learning (for a review, see Cohen and Newsome, 2004). The one caveat to this approach is that its functional specificity relies on a certain level of organization in the underlying neural architecture. Specifically, the technique has proved most successful when applied in brain areas or structures in which neurons with similar response properties are spatially clustered, presumably allowing for microstimulation to inject a coherent signal into the underlying neural circuitry.

Ultimately, more powerful circuit manipulation techniques may arise from promising developments in optogenetics (for a review, see Zhang *et al.*, 2007). Using genetic manipulations, small light-activated channel proteins are inserted into a circuit of choice. The channels can be either excitatory or inhibitory, and they can be activated and inactivated with excellent temporal precision simply by shining light on the neurons containing the channels, either via direct illumination for surface structures or via fiber-optic probes for deep structures. These techniques are potentially extremely powerful, and are being developed rapidly in species that are particularly amenable to genetic manipulation (such as mice and flies). Introduction of the channel genes into non-human primates via viral vectors appears feasible, and several labs are initiating work in this area. Application to humans, though a seemingly distant prospect at the moment, might arrive sooner than we think. This fact emphasizes how important it is that human and animal studies proceed in parallel. Work in non-human primates is particularly important in this regard, as these animals can perform many of the same complex decision-making tasks as humans. Fortunately, this parallel development has been a feature of the field of neuroeconomics since its inception, and initial signs suggest that a good deal of the basic neural machinery of decision making and valuation may be shared between species.

In neuroeconomics, we have yet to test causal links between particular neural circuits and behavioral choice. It is exciting to imagine a day in the near future when microstimulation, optogenetics, or other techniques might be used to test the hypotheses that are beginning to emerge about the functional roles of various structures. With so many of the fundamental discoveries in this field still on the horizon, and the tools with which to explore them finally becoming available, we anticipate that studying "choice" will continue to be worth the "trouble" for a considerable time to come.

References

Barraclough, D.J., Conroy, M.L., and Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* 7, 404–410.

Bechara, A., Damasio, H., and Damasio, A.R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex* 10, 295–307.

Burnham, K.P. and Anderson, D.R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York, NY: Springer-Verlag.

Cohen, M.R. and Newsome, W.T. (2004). What electrical microstimulation has revealed about the neural basis of cognition. *Curr. Opin. Neurobiol.* 14, 169–177.

Corrado, G. S. (2007). “Leveraging behavioral models to reveal the neural representation of value.” *Doctoral Dissertation*, Stanford University.

Corrado, G. and Doya, K. (2007). Understanding neural coding through the model-based analysis of decision making. *J. Neurosci.* 27, 8178–8180.

Corrado, G.S., Sugrue, L.P., Seung, H.S., and Newsome, W.T. (2005). Linear-nonlinear-Poisson models of primate choice dynamics. *J. Exp. Anal. Behav.* 84, 581–617.

Daw, N.D., O’Doherty, J.P., Dayan, P. *et al*. (2006). Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.

Gold, J.I. and Shadlen, M.N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574.

Hallett, M. (2007). Transcranial magnetic stimulation: A primer. *Neuron* 55, 187–199.

Hampton, A.N., Bossaerts, P., and O’Doherty, J.P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J. Neurosci.* 26, 7836–8360.

Hastie, T., Tibshirani, R., and Friedman, J. (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer.

Herrnstein, R.J. (1961). Relative and absolute strength of responses as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* 4, 267–272.

Herrnstein, R.J. and Prelec, D. (1991). Melioration: a theory of distributed choice. *J. Econ. Persp.* 5, 137–156.

Hikosaka, O. and Wurtz, R.H. (1985). Modification of saccadic eye movements by gaba-related substances. I. Effect of muscimol and bicuculline in monkey superior colliculus. *J. Neurophysiol.* 53, 266–291.

Holt, C.A. and Laury, S.K. (2002). Risk aversion and incentive effects. *Am. Econ. Rev.* 92, 1644–1655.

Kennerley, S.W., Walton, M.E., Behrens, T.E. *et al*. (2006). Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* 9, 940–947.

Knutson, B., Westdorp, A., Kaiser, E., and Hommer, D. (2000). fMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage* 12, 20–27.

Kringelbach, M.L., O’Doherty, J., Rolls, E.T., and Andrews, C. (2003). Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cerebral Cortex* 13, 1064–1071.

Lau, B. and Glimcher, P.W. (2005). Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* 84, 555–579.

McClure, S.M., Berns, G.S., and Montague, P.R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38, 339–346.

McPeek, R.M. and Keller, E.L. (2004). Deficits in saccade target selection after inactivation of superior colliculus. *Nat. Neurosci.* 7, 757–763.

Nash, J. (1950). equilibrium points in n-person games. *PNAS* 36, 48–49.

Ochsner, K.N., Knierim, K., Ludlow, D.H. *et al*. (2004). Reflecting upon feelings: an fMRI study of neural systems supporting the attribution of emotion to self and other. *J. Cogn. Neurosci.* 16, 1746–1772.

O’Doherty, J.P., Dayan, P., Friston, K. *et al*. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron* 38, 329–337.

O’oherty, J.P., Hampton, A., and Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Ann. NY. Acad. Sci.* 1104, 35–53.

Padoa-Schioppa, C. and Assad, J.A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.

Paton, J.J., Belova, M.A., Morrison, S.E., and Salzman, C.D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439, 865–870.

Plassmann, H., O’Doherty, J., and Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27, 9984–9988.

Reale, R.A. and Imig, T.J. (1980). Tonotopic organization in auditory cortex of the cat. *J. Comp. Neurol.* 192, 265–291.

Rolls, E.T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex* 10, 284–294.

Romo, R. and Salinas, E. (2003). Flutter discrimination: neural codes, perception, memory and decision making. *Nat. Rev. Neurosci.* 4, 203–218.

Ross, L. and Nisbett, R.E. (1991). *The Person and the Situation*. Philadelphia, PA: Temple University Press.

Salzman, C.D., Britten, K.H., and Newsome, W.T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. *Nature* 346, 174–177.

Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.

Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2005). Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* 6, 363–375.

Sutton, R.S. and Barto, A.G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Turing, A. (1950). Computing machinery and intelligence. *Mind* 236, 433–460.

van der Heijden, F., Duin, R.P.W., de Ridder, D., and Tax, D.M.J. (2004). *Classification, Parameter Estimation and State Estimation: An Engineering Approach Using MATLAB*. Chichester: Wiley.

Zhang, F., Aravanis, A.M., Adamantidis, A. *et al*. (2007). Circuitbreakers: optical technologies for probing neural signals and systems. *Nat. Rev. Neurosci.* 8, 577–581.

CHAPTER

31

Mechanisms for Stochastic Decision Making in the Primate Frontal Cortex: Single-neuron Recording and Circuit Modeling

Daeyeol Lee and Xiao-Jing Wang

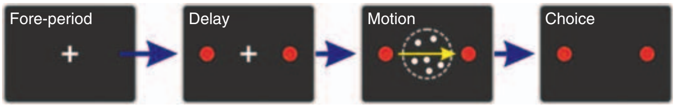
OUTLINE

INTRODUCTION

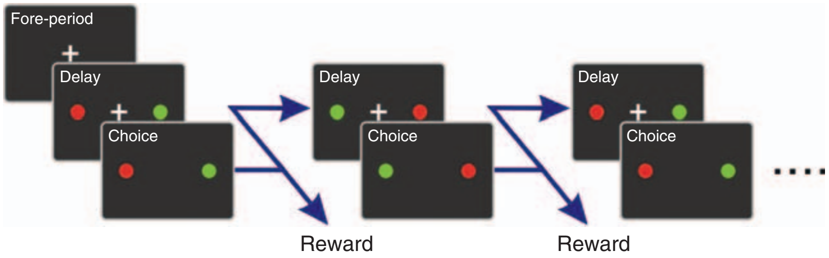
Decision Making refers to the process of evaluating the desirabilities of alternative choices and selecting a particular option. Thus, many brain functions can be characterized as decision making. For instance, animals seldom behave solely based on their intrinsic metabolic needs, as sensory information about the environment must be taken into account in determining which action the animal should produce to obtain the most desirable outcomes. Sensory inputs are always noisy, and perceptual analysis of these inputs reduces the uncertainty about the nature of sensory stimuli encountered in the animal's environment in order to categorize stimuli and select the most likely interpretation of the world. Once the relevant stimuli in the animal's environment are properly interpreted, it is then necessary to evaluate the desirability of the outcome expected from each of the alternative behavioral plans. Finally, even when the behavioral goal is set, a particular spatio-temporal sequence of muscle activation must be chosen to carry out the desired action.

Theoretical analyses of steady-state choice behavior are often formulated based on the principles of optimality and equilibrium. For example, game theory seeks to describe optimal decision-making strategies for multiple decision makers trying to maximize their own self-interests through a variety of complex social interactions (von Neumann and Morgenstern, 1944; Nash, 1950; see also Chapter 5 for an introduction to game theory). However, such approaches are limited when the decision makers do not possess all the information about the environment, or when the environment changes frequently. In such cases, the decision makers may need to improve their decision-making strategies incrementally by trial and error. This is often referred to as *melioration* (Herrnstein *et al*., 1997). Similarly, models based on reinforcement learning (Sutton and Barto, 1998) have been developed to account for how various choice behaviors change dynamically.

Neuroeconomics applies a variety of techniques to investigate the biological processes responsible for decision making. Among them, physiological techniques that probe the activity of individual neurons at millisecond temporal resolution are generally applicable only to animals, due to their invasive nature. In particular, non-human primates can be trained to perform relatively complex decision-making tasks, and their choice behaviors display many features similar to human choice behaviors. Since the brains of these two species display many structural and functional similarities, the results of single-neuron recording experiments conducted in non-human primates can provide valuable insights into the basic neural circuit mechanisms of decision making in humans. In recent years, electrophysiological studies in behaving monkeys have begun to uncover single-neuron activity that is correlated with specific aspects of perceptual decision making and reward-dependent choice behavior. For instance, Shadlen and his colleagues used a perceptual decision-making task based on random dot motion direction discrimination (Figure 31.1a), and found that activity of individual neurons in the primate posterior parietal cortex reflects the process of accumulating evidence (Roitman and Shadlen, 2002). Others (Surgue *et al*., 2004; Lau and Glimcher, 2005) examined the choice behavior of monkeys during a decision-making task based on concurrent schedules of reinforcement (Figure 31.1b), and found that their choice behavior largely conformed to the predictions of the matching law (Herrnstein *et al*., 1997). Furthermore, the activity of neurons in the posterior parietal cortex encoded the rate of reward or utility expected from a particular choice (Dorris and Glimcher, 2004; Sugrue *et al*., 2004). In contrast to perceptual discrimination tasks and concurrent schedules of reinforcement, competitive games involve interactions among multiple decision agents. It has been shown that monkeys are capable of producing stochastic choice behaviors that are nearly optimal for such competitive games (Dorris and Glimcher,



(a) Random-dot motion direction discrimination task



(b) Concurrent schedules of reinforcement

FIGURE 31.1 Decision-Making tasks used in monkeys. (a) A random-dot motion discrimination task. When the animal fixates a central target, two peripheral choice targets are presented. Then, random-dot motion stimuli are presented, and the animal is required to shift its gaze towards one of the choice targets according to the direction of random-dot stimuli. (b) During a decision-making task based on concurrent reinforcement schedules, each target is baited with a particular probability (variable rate) or after a time interval sampled from a particular distribution (variable interval).

2004; Lee *et al.*, 2004, 2005). During such tasks, neurons in the dorsolateral prefrontal cortex and the anterior cingulate cortex exhibited firing activity that reflected history of past choice and rewards (Barraclough *et al.*, 2004; Seo and Lee, 2007; Seo *et al.*, 2007).

These experiments have spurred theoretical work that uses mathematical approaches to illuminate experimental observations. For instance, accumulator models have been widely applied to perceptual decision making (Smith and Ratcliff, 2004; Gold and Shadlen, 2007; see also Chapter 4). Reward-based choice behavior has been described by reinforcement learning models (Sutton and Barto, 1998; see also Chapter 22). We will briefly summarize these models in relationship to our neurophysiological recording experiments in non-human primates. Our focus, however, will be neural circuit modeling, which attempts to go one step further and explore how decision behavior and correlated neural activity can be explained by the underlying circuit mechanisms. For example, what are the neural circuit substrates for time integration of sensory evidence about alternative choices and for action selection? Is valuation of actions instantiated by neurons or synapses, and how does a neural circuit make dynamic decisions adaptively over time? What are the sources of stochasticity in the brain that underlie random choice behavior? We will introduce neural circuit models and illustrate their applications to a perceptual discrimination task (Wang, 2002; Machens *et al.*, 2005; Lo and Wang, 2006; Wong and Wang, 2006; Miller and Wang, 2006), a foraging task based on concurrent schedules of reinforcement (Soltani and Wang, 2006), and a matching pennies game task (Soltani *et al.*, 2006). These computational studies showed that a general neural circuit model can reasonably account for salient behavioral and electrophysiological data from diverse decision tasks, suggesting common building blocks of decision-making circuitry that may be duplicated throughout different stages of sensori-motor transformation in the primate brain.

GAME-THEORETIC STUDIES OF DECISION MAKING IN NON-HUMAN PRIMATES

Game Theory and Learning in Social Decision Making

When each of the alternative actions produces a particular outcome without any uncertainty, optimal decision making consists simply of choosing the action that produces the most desirable outcome. When there is uncertainty about the outcomes expected from various actions, the animal's choice should be influenced by the likelihood of desirable outcomes expected from each action. A large number of economic theories, such as the expected utility theory (von Neumann and Morgenstern, 1944) and prospect theory (Kahneman and Tversky, 1979), have been proposed to account for such decision making under uncertainty or risk. In reality, however, the environment changes constantly, and this frequently alters the likelihood of various outcomes resulting from different actions. Consequently, optimality is rarely achieved, and typically subjects can only approximate optimal decision strategies by learning through experience (Sutton and Barto, 1998).

The complexity and hence difficulty of such learning would depend on the nature of dynamic changes in the animal's environment, which can occur for a number of reasons. Some are cyclical, such as seasonal changes, and others are predictable changes resulting from the animal's own actions, such as the depletion of food and other resources. Animals living in social groups face even more difficult challenges, because competitive interactions with other animals tend to make it quite difficult to predict the outcomes resulting from their own actions. Nevertheless, decision making in such social settings provides a unique opportunity to test various theories about the behavioral dynamics and underlying mechanisms of decision making.

One way to tackle mathematically the problems of decision making in a social context is formulated by game theory (von Neumann and Morgenstern, 1944; see also Chapter 5 of this volume). In game theory, a game is specified by a particular number of decision makers or players, a list of alternative choices available to each player, and the payoff matrix that assigns a particular outcome or payoff to each player according to the combination of actions chosen by all players. In other words, the payoff given to a player does not depend simply on that player's own action, but on the actions of all players in the game. In addition, a strategy for a given player is defined as a probability distribution over a set of available choices. A pure strategy refers to choosing a particular action exclusively, whereas a mixed strategy refers to a case in which multiple actions are chosen with positive probabilities. One of the predictions from game theory is that a set of players trying to maximize their self-interests would converge onto a set of strategies known as Nash Equilibrium, which is defined as a set of strategies for all players that cannot be changed by any individual player to increase his payoff. A game is called a mixed-strategy game when its equilibrium strategy is mixed. In addition, a game is referred to as zero-sum when the sum of the payoffs for all players is always zero, so that someone's gain necessarily means someone else's loss. In the following sections, we describe the results from behavioral experiments to illustrate that non-human primates can approximate the mixed equilibrium strategies through iterative learning algorithms in competitive zero-sum games with two and three alternative actions. These correspond to the familiar games of matching pennies and Rock–Paper–Scissors, respectively.

Choice Behavior During a Matching Pennies Game

To test whether and how monkeys approximated optimal decision-making strategies in simple competitive games through experience, previous studies have examined the choice behavior of monkeys in a computer-simulated binary zero-sum game, commonly referred to as matching pennies (Lee *et al*., 2004; Figure 31.2a). In this game, each of two players chooses from two alternative options, and one of the players wins if his choices match and loses otherwise. During the experiment, a monkey was required to begin each trial by fixating a small yellow square presented in the center of a computer screen. Shortly thereafter, two identical green disks were presented along the horizontal meridian, and the animal was required to shift its gaze towards one of the targets when the central fixation target was extinguished. The computer opponent also chose its target according to a pre-specified algorithm described below, and the animal was rewarded only when it chose the same target as the computer (Figure 31.2b).

When two rational players participate in the matching pennies game, the Nash Equilibrium is for each player to choose the two targets with equal probabilities and independently across successive trials. Any other strategy could be exploited by the opponent. In the monkey experiment, the strategy of the computer opponent was systematically manipulated to determine how the animal's choice behavior would be affected by the strategy of its opponent. Initially, the computer opponent chose the two targets with the same probabilities, regardless of the animal's choices. This was referred to as algorithm 0, and corresponded to the Nash Equilibrium strategy pursued unilaterally without regard to the opponent's behavior. In this case, the animal's expected payoff would be fixed regardless of how it chose its target. Therefore, it was not surprising that all three monkeys tested with algorithm 0 displayed a strong bias to choose one of the two targets more frequently. Overall, the monkeys C, E, and F chose the right-hand target in 70.0%, 90.2%, and 33.2% of the trials, respectively. In the next stage of the experiment, the computer opponent applied a set of statistical tests to determine whether the animal's choice was randomly divided between the two targets. If not, the computer decreased its probability of choosing a particular target as the animal chose the same target more frequently. This was referred to as algorithm 1. Importantly, this algorithm did not examine the animal's reward history, and therefore was not sensitive to any bias that the animal might show in using the information about its previous rewards to determine its future choices. When tested with algorithm 1, the animals chose the two targets more or less equally frequently. Overall, during algorithm 1, the monkeys C, E, and F chose the right-hand target

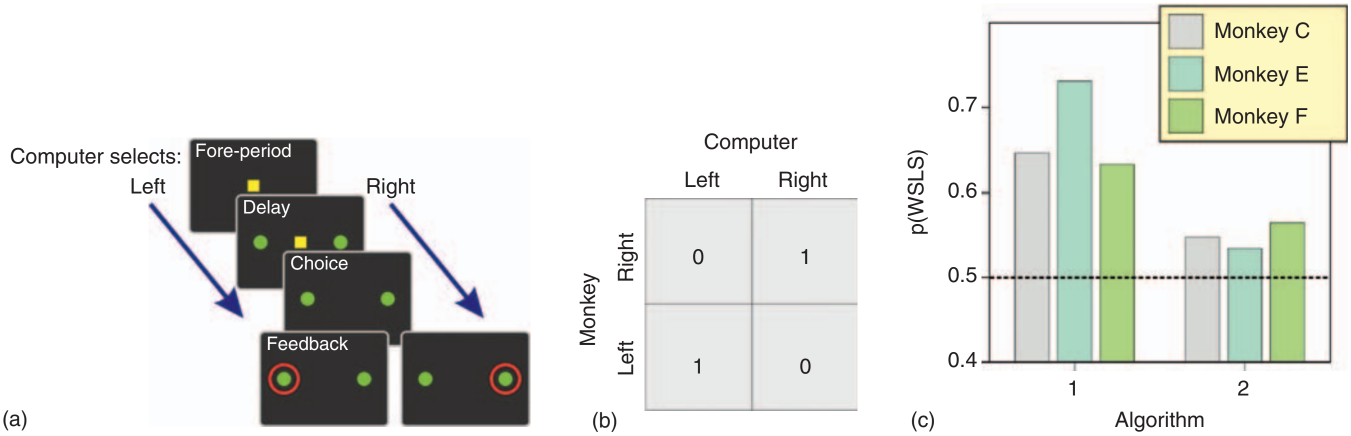


FIGURE 31.2 Monkey's choice behavior during the matching pennies game.(a) During this task, the animal made a saccadic eye movement towards one of the two peripheral targets to indicate its choice and was rewarded only when it chose the same target as the computer opponent. (b) The payoff matrix for the matching pennies game. (c) The probability that the animal would choose its target according to the so-called win–stay–lose–switch strategy during the matching pennies game against the computer opponent programmed with two different algorithms (1 and 2).

in 48.9%, 51.1%, and 49.0% of the trials, respectively. In addition, the animal's successive choices were relatively independent, and as a result, the animal's overall reward rate was close to the optimal value of 0.5 (Lee *et al*., 2004). Interestingly, the animals were more likely to choose the same target in the next trial if it was rewarded in a given trial (win–stay) and switch to the other target otherwise (lose–switch). Such strategies were not penalized during the period of algorithm 1, since the information about the animal's reward history was not utilized by the computer opponent. All three animals chose their targets according to this so-called win–stay–lose–switch (WSLS) strategy substantially more than in 50% of the trials (Figure 31.2c).

In reinforcement learning models (Sutton and Barto, 1998; Camerer, 2003), the animal's choice is modeled by a set of value functions that are adjusted according to the outcome of the animal's choice. To test whether the animal's choice during the matching pennies game was accounted for by a reinforcement learning model, the value function at trial t for a given target x, $V_t(x)$, was updated after each trial according to the following (Lee *et al*., 2004):

$$V_{t+1}(x) = \alpha\ V_t(x) + \Delta_t(x) \qquad (31.1)$$

where $x = L$ and R for the right-hand and left-hand targets, respectively, α is a decay factor, and $\Delta_t(x)$ reflects a change in the value function that depends on the outcome of a choice. It was assumed that if the animal chose the target x, the value function for x was adjusted according to whether the animal was rewarded or not. Namely, when the animal selected the target x in trial t, $\Delta_t(x) = \Delta_{\text{rew}}$ if the animal was rewarded, and $\Delta_t(x) = \Delta_{\text{unrew}}$ otherwise. For the target not chosen by the animal, $\Delta_t(x) = 0$. The probability that the animal would select the right-hand target was then determined by the following softmax rule:

$$p_t(\text{R}) = \exp V_t(\text{R})/\{\exp V_t(\text{R}) + \exp V_t(\text{L})\} \qquad (31.2)$$

The model parameters, α, Δ_{rew}, and Δ_{unrew} were estimated using a maximum likelihood procedure (Pawitan, 2001). This reinforcement learning model was compared with an alternative model based on the WSLS strategy. This WSLS model assumed that the animal chooses its target in each trial according to the WSLS strategy with some probability, p_{WSLS}. For example, if the animal was rewarded for choosing a particular target in the previous trial, this model assumes that the animal would choose the same target in the current trial with the probability of p_{WSLS}. If the animal was not rewarded in the previous trial, it would switch to the other target with the same probability. Whether the animal's choice behavior was better accounted for by the reinforcement learning model or the WSLS model was determined by the Bayesian Information Criterion (BIC), defined as

$$\text{BIC} = -\log 2 \log \text{L} + \text{k} \log \text{N} \qquad (31.3)$$

where L denotes the likelihood of the model, k the number of model parameters (k = 1 and 3 for the WSLS model and reinforcement learning model, respectively), and N the number of trials (Hastie *et al*., 2001). For each session, the model that minimized the BIC was chosen as the best model. The results of model-fitting showed that for the reinforcement learning model applied to the choice data obtained with algorithm 1, the value functions for a given target tended to increase (decrease) when the animal was (not) rewarded for choosing the same target (Figure 31.3). In addition, the reinforcement learning model performed better than the WSLS model in 56.0% of the sessions (70/125 sessions). Therefore, for algorithm 1, the performance of these two models was comparable, suggesting that the animal's choice was largely determined by the WSLS strategy.

The same animals were then further tested against a third computer opponent which tested not only whether the animal's choice sequences were random, but also whether the animal's choice was independent of its previous choice outcomes (reward vs no reward). This was referred to as algorithm 2 (Lee *et al*., 2004). Compared to the choice behavior during algorithm 1, the choice behavior resulting from algorithm 2 was dramatically different in that now the tendency to use the WSLS strategy was significantly reduced in all animals (Figure 31.2c). When the choice behavior was analyzed with the reinforcement learning model, it was found that the trial-to-trial changes in the value functions were much smaller compared to the results obtained in algorithm 1 (Figire 31.3). In addition, the reinforcement learning algorithm now accounted for the choice behavior better than the WSLS model in 63.9% of the sessions (62/97 sessions). Therefore, the animals reduced their reliance on the WSLS strategy when it was no longer profitable.

Choice Behavior During a Rock–Paper–Scissors Game

Two monkeys were tested in a computerized Rock–Paper–Scissors game (Figure 31.4a). During this game, the Nash Equilibrium is to choose randomly each of the three options with the probability of 1/3. Any other strategy can be exploited by the opponent.

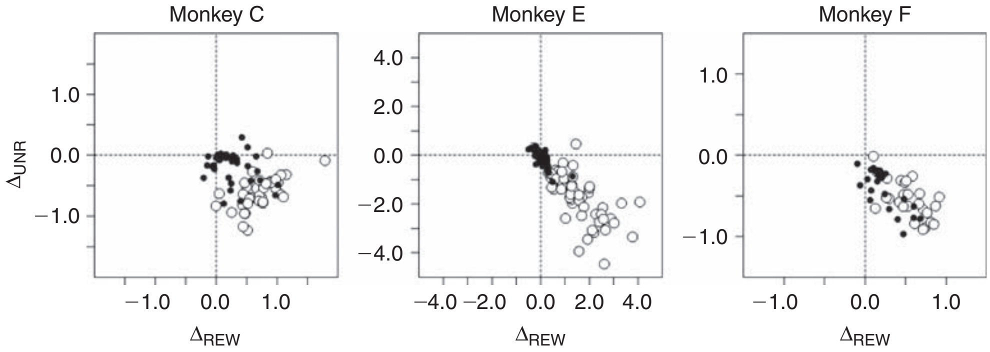
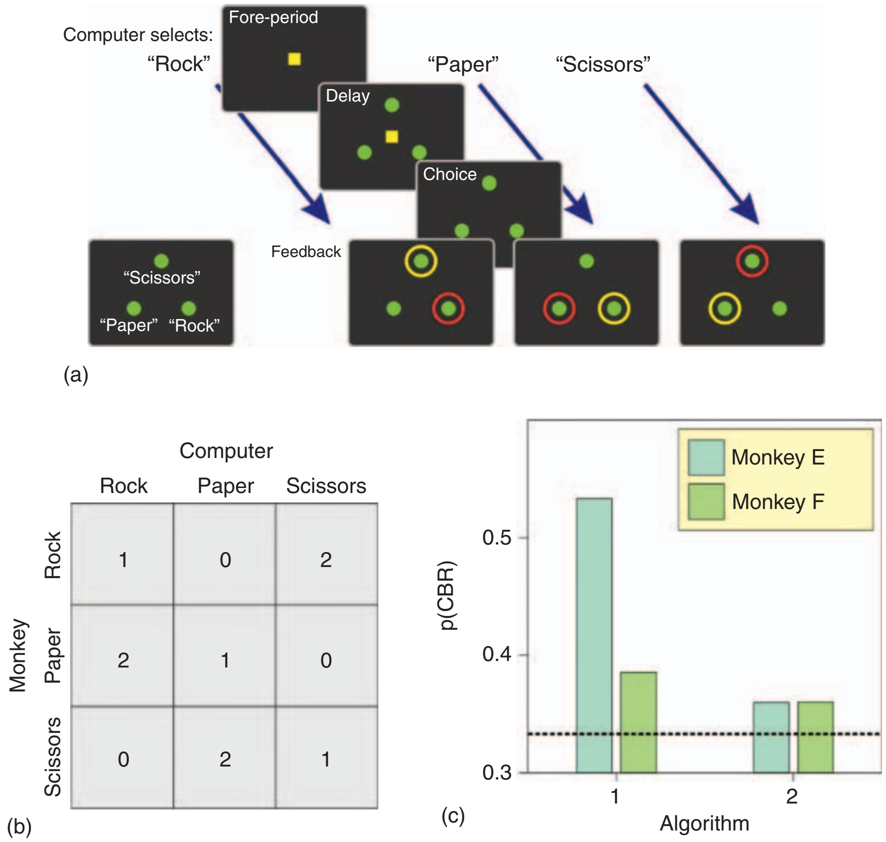


FIGURE 31.3 Parameters in the reinforcement learning model fit to the choice behavior of three different monkeys tested in the matching pennies games. The large empty circles and small filled circles correspond to the parameters obtained for individual sessions tested with algorithms 1 and 2, respectively.



		Computer		
		Rock	Paper	Scissors
Monkey	Rock	1	0	2
	Paper	2	1	0
	Scissors	0	2	1

FIGURE 31.4 Monkey's choice behavior during the rock–paper–scissors game. (a) During this task, the animal chooses among three alternative targets, and receives different payoffs according to the rules of the rock–paper–scissors game. (b) The payoff matrix for the rock–paper–scissors game. (c) The probability that the animal would choose its target in the rock–paper–scissors game corresponding to the so-called Cournot best response against the opponent using two different algorithms.

Similar to the matching pennies game, each trial began when the animal fixated a small yellow square at the center of the computer screen. Then, three identical green disks were presented, with their spatial positions arbitrarily designated as "Rock," "Paper," and "Scissors." As in the matching pennies task, the computer chose its target according to one of three different algorithms, and the animal was rewarded according to the payoff matrix of the Rock–Paper–Scissors game (Figure 31.4b). Namely, the animal was rewarded by a drop of fruit juice when it chose the same target as the computer opponent, and by two drops of juice when its choice beat the computer's (e.g., when the animal and the computer chose

Rock and Scissors, respectively). The animal was not rewarded when its choice was beaten by the computer's choice.

Similar to the results obtained from the matching pennies task, each animal displayed an idiosyncratic bias to choose a particular target, while the computer opponent unilaterally complied with the Nash Equilibrium by choosing the three targets with the same probabilities regardless of the animal's choices (Lee *et al.*, 2005; algorithm 0). Furthermore, as in the matching pennies task, when the computer opponent predicted the animal's choice and behaved competitively using only the information about the animal's choice sequences (algorithm 1), the animal started choosing the three targets with more or less equal probabilities. Nevertheless, there was a significant bias to choose the target that would have beaten, and therefore was the best response to, the computer's choice in the previous trial. For example, if the animal chose Rock in a given trial, it tended to choose Scissors in the next trial. This is referred to as the *Cournot best response* (Lee *et al.*, 2005), which becomes the WSLS strategy when there are only two alternative choices. In the final stage of the experiment (algorithm 2), the computer opponent utilized information about both the animal's choice sequences and their outcomes. Therefore, the computer could now detect whether there was any bias for the animal to rely on such strategies as the Cournot best response. During algorithm 2, the probability of using the Cournot best response was significantly reduced in both animals. Nevertheless, both animals chose their targets corresponding to the Cournot best response significantly more frequently than expected by chance (36.0% for both animals; Figure 31.4c). These results were largely consistent with the predictions of a reinforcement learning algorithm (Lee *et al.*, 2005).

NEUROPHYSIOLOGICAL STUDIES OF DECISION MAKING IN COMPETITIVE GAMES

The results described in the previous section suggest that monkeys might approximate equilibrium strategies in competitive games using reinforcement learning algorithms. Recently, a relatively large number of studies have identified neural signals related to the key components of reinforcement learning in multiple brain areas. For example, signals related to the discrepancy between the predicted and actual rewards, commonly referred to as reward-prediction error, have been found in the anterior cingulate cortex (Matsumoto *et al.*, 2007; Seo and Lee, 2007) as well as the midbrain dopamine neurons (Schultz, 1998). In addition, signals resembling value functions have been identified in various cortical areas and the basal ganglia (see the other chapters in Part 5 of this volume). Thus, many of these cortical and subcortical areas might also be involved in updating the animal's decision-making strategy during competitive games. The studies described below tested this in the dorsolateral prefrontal cortex and the anterior cingulate cortex (Barraclough *et al.*, 2004; Seo and Lee, 2007; Seo *et al.*, 2007).

Role of the Dorsolateral Prefrontal Cortex in Decision Making

The primate dorsolateral prefrontal cortex (DLPFC) is anatomically connected with the multiple cortical areas involved in the processing of sensory, motor, and affective signals (Barbas and Pandya, 1987; Vogt and Pandya, 1987; Barbas, 1995; Carmichael and Price, 1996; Fuster, 1997; Luppino *et al.*, 2003). Consistent with such diverse sources of its inputs, neurons in the DLPFC display modulations in their activity that reflect various aspects of sensory stimuli, motor responses, and mnemonic processes (Chafee and Goldman-Rakic, 1998; Constantinidis *et al.*, 2001; Amemori and Sawaguchi, 2006). In addition, many neurons in the DLPFC change their activity according to the nature of the reward expected from a particular state of the environment (Watanabe, 1996; Leon and Shadlen, 1999; Kobayashi *et al.*, 2002). These findings suggest that the DLPFC might play an important role in decision making by encoding a particular state of the environment and the desirability of the outcome expected from such a state (Lee *et al.*, 2007).

Whether the neurons in the DLPFC encode signals necessary to optimize the animal's decision-making strategy was tested during the matching pennies task. The results from these experiments showed that the DLPFC neurons encode a multitude of signals that can contribute to different stages of decision making (Barraclough *et al.*, 2004; Seo *et al.*, 2007). First, some neurons in the DLPFC modulated their activity according to the value functions of the two alternative targets, and therefore might be involved in computing or encoding the decision variable necessary to determine the animal's choice. Second, activity of many neurons in the DLPFC encoded not only the upcoming choice of the animal, but also the animal's choices in the previous trials (Figure 31.5, top; Figure 31.6, top). In real life, the consequence of a particular action might be revealed after some temporal delay and often after several intervening actions. This leads to the problem

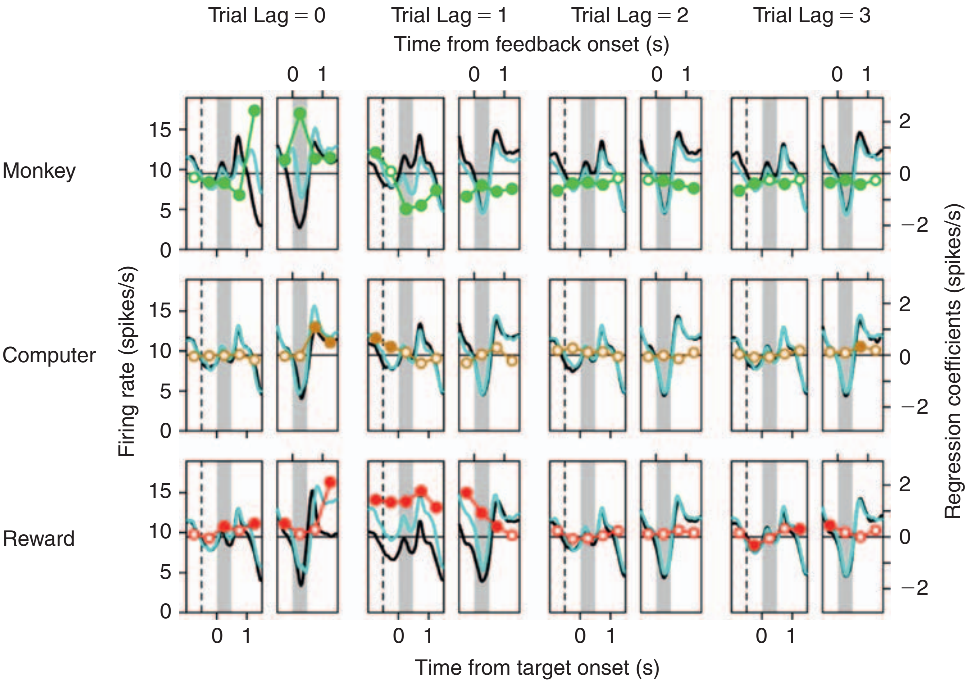


FIGURE 31.5 An example neuron recorded in the dorsolateral prefrontal cortex of a monkey playing the matching pennies game against a computer opponent. Each panel shows the average activity of the neuron aligned to the onset of the targets or the feedback stimulus, separately according to the animal's choice (top), the choice of the computer opponent (middle), or the reward (bottom) in the current (Trial lag = 0) or previous trials (Trial lag = 1 to 3). Black and blue lines indicate whether the animal (top) or the computer (middle) selected the left-ward and right-ward target, or whether the animal was rewarded (black) or not (blue). Filled and empty circles correspond to the regression coefficients quantifying the effect of each variable in different time bins and also indicate whether each coefficient was statistically significant (filled; *t*-test, $P < 0.05$) or not (empty). Dotted vertical lines indicate the time when the animal fixated the central target. The gray background in the left panels corresponds to the cue period during which the animal was required to maintain its fixation on the central target, whereas the gray background in the right panels corresponds to the feedback period.

of temporal credit assignment, since it may be important to identify correctly the action that was responsible for a particular outcome. Memory signals related to the animal's previous actions, such as those found in the DLPFC, are referred to as eligibility trace, and may play an important role in solving such problems. Third, some neurons in the DLPFC modulated their activity according to the previous choices of the computer opponent (Figure 31.5, middle; Figure 31.6, middle). During the matching pennies game, the average reward expected from choosing a particular target is directly related to the probability that the same target will be chosen by the computer opponent. Therefore, the signals related to the computer's previous choices provide the information necessary to estimate the likelihood of reward from a particular target. Finally, activity of the DLPFC neurons during the feedback period was often influenced by the outcomes of the animal's previous choices, namely, whether the animal was rewarded or not in the previous trials (Figure 31.5, bottom; Figure 31.6, bottom). This suggests that the activity in the DLPFC might be influenced by the animal's reward history and therefore by the context in which a particular reward is delivered or omitted. Overall, these results suggest that the DLPFC might be closely involved in updating the animal's decision-making strategies as well as determining the animal's choice in a given trial.

Role of the Anterior Cingulate Cortex in Decision Making

The primate anterior cingulate cortex is closely linked to the lateral prefrontal cortex and the orbitofrontal cortex (Vogt and Pandya, 1987; Barbas and Pandya, 1989; Carmichael and Price, 1996; Luppino *et al.*, 2003). Similar to the lateral prefrontal cortex, the anterior cingulate cortex is also connected to cortical areas associated with motor functions, such as the supplementary motor area and supplementary eye-field (Wang *et al.*, 2001; Luppino *et al.*, 2003). Compared to the lateral prefrontal cortex, however, the anterior cingulate cortex is more densely connected with the subcortical structures involved in emotion and reward processing, such as the amygdala (Barbas

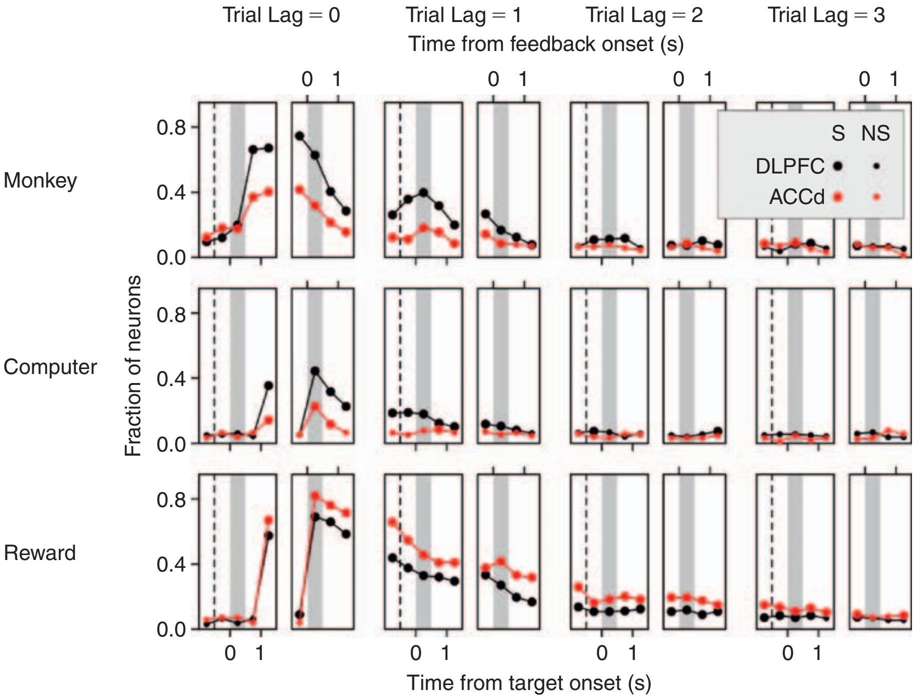


FIGURE 31.6 Fraction of neurons in the dorsolateral prefrontal cortex (black) and the dorsal anterior cingulate cortex (red) that significantly modulated their activity at a particular time bin according to the animal's choice (top), the choice of the computer opponent (middle) or the reward received by the animal (bottom) in the current (Trial lag = 0) or previous (Trial lag = 1 to 3) trials. Large symbols indicate that the number of neurons with significant effects was larger than expected by chance (binomial test, $P < 0.05$).

and De Olmos, 1990; Ghashghaei *et al.*, 2007) and the ventral tegmental area (Williams and Goldman-Rakic, 1993; Lewis *et al.*, 2001). Accordingly, neurons in the anterior cingulate cortex often modulate their activity according to the expected and actual reward (Niki and Watanabe, 1979; Amiez *et al.*, 2005, 2006), but neural signals reflecting the information about specific actions are less frequently observed in the anterior cingulate cortex than in the lateral prefrontal cortex (Hoshi *et al.*, 2005; Matsumoto *et al.*, 2007).

Consistent with these findings, the activity of neurons in the anterior cingulate cortex was commonly modulated by the outcomes of the animal's choices during the matching pennies game. For example, many ACC neurons increased or decreased their activity when the animal received a feedback signal indicating that it would be rewarded in a given trial. Such activity related to the choice outcome or reward was often maintained across multiple trials, and, as a result, the reward-related activity in the ACC was often influenced by the animal's reward history (Seo and Lee, 2007). In addition, the time-course of such activity related to the outcome of the animal's previous choice varied substantially across different neurons. In some neurons, activity was influenced similarly by the animal's rewards in the current and previous trials, suggesting that the activity of such neurons might signal the local rate of reward. Such signals might be used as a reference point against which the value of the reward in a current trial is evaluated. In contrast, the activity of other neurons was modulated antagonistically by the rewards in the current and previous trials. For example, a particular neuron might increase its activity when the animal is rewarded in the current trial, but its activity might be decreased if the animal was rewarded in the previous trial. Such neurons would then show their maximal or minimal activity when the animal was rewarded after losing a reward in the previous trial. Analogous to the activity of dopamine neurons that encode reward-prediction error (Schultz, 1998), some neurons in the ACC might therefore encode the deviation of the reward received by the animal in the current trial relative to the reward expected from the outcomes in the previous trial. To test this further, the value functions of the two alternative targets as well as the reward-prediction error were estimated on a trial-by-trial basis using a standard reinforcement learning model, and these estimates were used to test whether the activity of each ACC neuron was correlated with the sum or difference of the value functions or with the reward-prediction error. The results showed that the activity of many ACC neurons was correlated with the sum of the value function, but relatively infrequently with the difference of value functions. Since reinforcement learning algorithms choose a particular action based on the difference in value functions, these results suggest that the ACC makes only a minor contribution to the process of action selection. In contrast, the activity of many neurons in the ACC was strongly correlated with the reward-prediction error. Presumably, outcome-related activity in the ACC that is modulated by the animal's reward history might influence the process of updating the value function of the alternative targets and therefore the animal's subsequent actions. Precisely how this is accomplished, however, is currently not known.

MODELS OF DECISION MAKING

Experiments using different types of decision behaviors have led to two broad classes of models for decision making. On the one hand, so-called sequential-sampling models describe information processing that unfolds in time and determine performance accuracy and reaction times in perceptual and memory tasks. One the other hand, game-theoretical models and reinforcement learning models account for dynamic choice behavior which is based on utility maximization and interplay with the environment or other decision agents. These models are important for quantitatively describing behavioral data and assessing theoretical ideas about the cognitive processes of decision making. To truly understand the biological basis of decision behavior, however, it is critical to construct realistic neural circuit models that allow us to explore how key computations of decision making are instantiated by cellular machineries and collective dynamics of neural networks in the brain. Here we summarize models of varying degrees of complexity, with a focus on recently advanced neural models of decision circuits. We show that both perceptual decisions and value-based choice behaviors can be described by a recurrent neural circuit model that is endowed with reward-dependent synaptic plasticity.

Drift-diffusion, Leaky Competing Accumulator, and Attractor Network Models

Sequential-sampling models are based on the intuitive idea that a decision is reached when stochastic accumulation of information about alternative choices reaches a particular threshold. For two-alternative forced choice tasks, a specific implementation is called the drift-diffusion model (Ratcliff, 1978; Smith and Ratcliff, 2004). In this model, an activity variable X represents the difference between the respective amounts of accumulated information about the two alternatives, say X_1 and X_2, $X = X_1 - X_2$. The dynamics of X is given by the drift diffusion equation,

$$dX/dt = \mu + w(t) \tag{31.4}$$

where μ is the drift rate and $w(t)$ is a white noise of zero mean and standard deviation σ. The drift rate μ represents the bias in favor of one of the two choices (and is zero if there is no net bias). For instance, in a random-dot motion direction discrimination task, μ is proportional to the strength of motion signal. This system is a perfect integrator of the input:

$$X(t) = \mu t + \int^{t} w(t')\, dt' \tag{31.5}$$

The integration process is terminated and the decision time is read out whenever $X(t)$ reaches a positive threshold θ (choice 1) or a negative threshold $-\theta$ (choice 2). If the drift rate μ is positive, then choice 1 is correct, while choice 2 is an error (Figure 31.7). Therefore, this type of models is commonly referred to as "ramping-to-threshold" model, with the average ramping slope given by μ.

The drift-diffusion model has been widely applied to fit behavioral data from perceptual and memory experiments (Ratcliff, 1978; Smith and Ratcliff, 2004). This model is the continuous-time equivalent of the discrete-time Sequential Probability Ratio Test (SPRT), which is the optimal procedure for making binary choices under uncertainty, in the sense that it

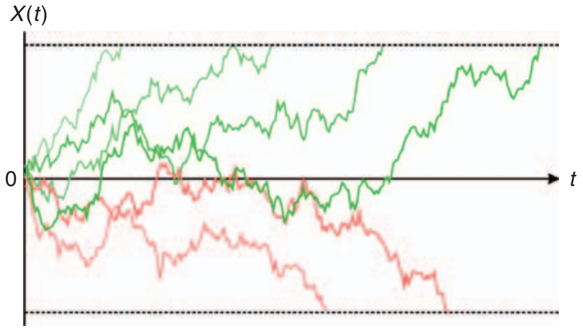


FIGURE 31.7 Drift diffusion model for time integration in reaction time tasks.The variable X(t) represents the difference in the activity at time t between two neural pools A and B. X(t) undergoes a biased random walk, and decision A (or B) is reached when X reaches a prescribed threshold θ (or $-\theta$). The time it takes to reach the threshold is read out as the response time. Six sample trials (four correct trials, green; two error trials, red) are shown.

minimizes the mean decision time among all tests for a given lower bound of error rate (Wald and Wolfowitz 1948; Bogacz *et al*., 2006).

Can a ramping-to-threshold mechanism be instantiated by neural circuits? One key issue here is to determine the biological basis of time integration. The drift-diffusion model is a perfect integrator, whereas neurons and synapses are leaky, with short time constants of tens of milliseconds (Kandel *et al*., 2000). Usher and McClelland (2001) extended the drift-diffusion model by incorporating a leak so that the integrator becomes forgetful with a decay time constant. Moreover, there is a competition between the two dynamical variables X_1 and X_2 through mutual inhibition. This leaky competitive accumulator model proved to fit many behavioral data equally as well as the drift-diffusion model, provided that the integration time is sufficiently long.

We have proposed that a long integration time can be realized biophysically in a decision neural network through recurrent excitation (Wang, 2002). Reverberating excitation represents a salient characteristic of cortical local circuits (Douglas and Martin, 2004). When this positive feedback is sufficiently strong, recurrent excitation in interplay with synaptic inhibition can create multiple stable states ("attractors"). Therefore, this class of models is referred to as "attractor network models." Such models have been initially proposed for working memory – namely, the brain's ability to actively hold information online in the absence of direct sensory stimulus. If a cortical network is endowed with a spontaneously active resting state and multiple memory states (for maintaining different items of information), transient inputs (such as a sensory cue) can switch the system from the resting state to one of the memory states, which is then self-sustained in the form of persistent neural activity. This idea has been explored using biophysically realistic neural circuit models (Wang, 2001; Amit and Mongillo, 2003). Recently, it has been realized that the same models can also implement decision-making computations (Wang, 2002; Machens *et al*., 2005; Miller and Wang, 2006; Wong and Wang, 2006). Interestingly, physiological studies in behaving nonhuman primates often reported neural activity correlated with decision making in cortical areas, such as the prefrontal cortex or the parietal cortex, that also exhibit mnemonic persistent activity during working memory. Hence, recurrent network (attractor) models can be viewed as a promising neuronal instantiation of the ramping-to-threshold model, and they serve as a framework to elucidate a common circuit mechanism for decision making and working memory in the brain.

What is Spiking Network Modeling?

Physiological experiments in behaving animals are critical to uncover neural signals correlated with specific aspects of decision making. Biophysically-based neural modeling can delineate circuit mechanisms that give rise to the observed neural signals, and identify key computational principles at the conceptual level. For certain questions about decision making, discussed below, it is important to capture neural firing of action potentials or spikes (electrical signals often described mathematically as point processes) through which neurons transmit information and communicate with each other. To this end, single cells should be described by a spiking neuron model, rather than a firing-rate model (as in the Usher-McClelland model). A popular choice is either the leaky integrate-and-fire model or the Hodgkin-Huxley model. Such a model is calibrated by physiological measurements, such as the membrane time constant and the input–output function (the spike firing rate as a function of the synaptic input current), which can be different for excitatory pyramidal cells and inhibitory interneurons. Furthermore, it is worth emphasizing that, in a biophysically-based model, synapses must be modeled accurately. Unlike connectionist models in which coupling between neurons is typically an instantaneous function of firing activity, synapses have their own rise-time and decay time constant, and exhibit summation properties. Synaptic dynamics turn out to be a crucial factor in determining the integration time of a neural circuit dedicated to decision making, as well as controlling the stability of a strongly recurrent network. Finally, "building blocks" (single cells and synapses) are used to construct a network endowed with a biologically plausible architecture. A commonly assumed circuit organization is the so called "Mexican-hat connectivity" – local excitation between neurons of similar selectivity combined with a more global inhibition (see Figure 31.8 for a schematic depiction of a simple two-neural pool network). Dynamic balance between synaptic excitation and inhibition is another feature of cortical microcircuits that has been increasingly recognized experimentally and incorporated in cortical network models.

Although this type of spiking network model may seem complicated compared with simple mathematical models such as the drift-diffusion model, it is actually minimal for quantitatively accounting for both behavioral data (determined by collective network dynamics) and spike train data of single neurons obtained from physiological experiments, and for exploring the synaptic circuit mechanisms required to explain these observations.

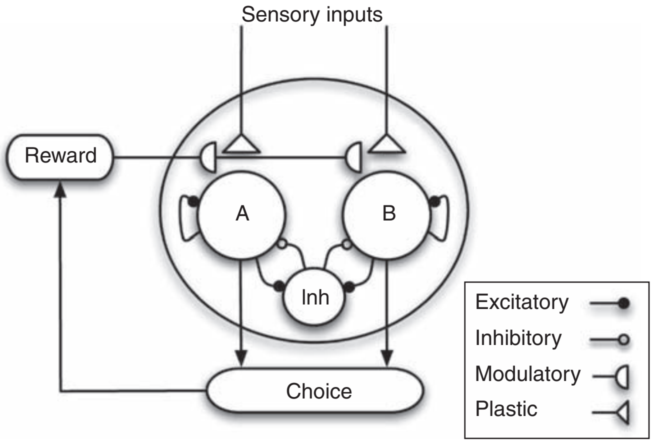


FIGURE 31.8 A simple model for decision making with two alternatives.There are two pyramidal cell groups, each of which is selective to one of the two choice options, A or B. Within each pyramidal neural group there are strong recurrent excitatory connections that can sustain persistent activity triggered by a transient preferred stimulus. The two neural groups compete through feedback inhibition from interneurons. Depending on the task design, one of the choices may be rewarded with some probability, whereas the other may not, in any given trial. The outcome signal (reward or not) is assumed to modulate Hebbian plasticity of input synapses c_A and c_B. Since the network's decision dynamics depends on c_A and c_B, altered synaptic strengths lead to adaptive choice behavior across trials.

A Recurrent Circuit Mechanism of Decision Making

A simple model for decisions with two alternative choices is schematically illustrated in Figure 31.8 (Wang, 2002). Two neural pools are selective for choice options (A or B), each consisting of a number of spiking neurons. The network connectivity exhibits salient features of a cortical circuit model: (1) direct excitatory connections are strong between pyramidal cells with similar selectivity (within each neural pool) and weak between those with entirely different selectivity (between the two neural pools); (2) there is a balance between synaptic excitation and inhibition; and (3) inhibitory interneurons induce competition between the two neural pools. As a result, interactions between the two neural pools are effectively inhibitory (Wong and Wang, 2006), similar to the Usher-McClelland model (Bogacz *et al.*, 2006). However, in contrast to the linear leaky competitor model, non-linear attractor dynamics that depend on reverberating excitation are critical to winner-takes-all competition.

Evidence for the two alternatives is provided by inputs I_A and I_B to the two neural pools. For instance, choice should be obvious if I_A is large and I_B is zero, whereas the weights of evidence for the two alternatives are equal if $I_A = I_B$. This is shown in Figure 31.9, where the input magnitudes are expressed in terms of the synaptic strengths c_A and c_B of the input pathway. Figures 31.9a and 31.9b display the time-course of network activity when c_A is larger than c_B in several simulation trials. In Figure 31.9a, the population firing rates r_A and r_B of the two neural pools first increase together, then start to diverge: r_A ramps up for hundreds of milliseconds, whereas r_B gradually decays away. Therefore, the decision outcome of the network is A, which is the correct choice. Figure 31.9b shows several error trials, in which the neural pool B rather than A wins the competition. Across many trials, the probability of choosing A (the correct response) is larger than 0.5 when evidence in favor of A is stronger than that for B (c_A is larger than c_B). Figure 31.9c shows that this probability is a sigmoid function of the difference $c_A - c_B$. Note that even when the inputs are equal ($c_A = c_B$), stochastic neural spiking dynamics still gives rise to winner-takes-all competition in each trial, with the choice probability at chance level (50%). Interestingly, this choice probability function is relatively insensitive to the absolute strengths c_A and c_B; therefore the network's decision performance can be described by a softmax function

$$P_A(c_A - c_B) = 1/(1 + \exp(-(c_A - c_B)/\sigma) \qquad (31.6)$$

where σ expresses the amount of stochasticity due to irregular spike firing in the network and also depends on other model properties, such as the firing rates of input neurons. Importantly, a softmax decision criterion is widely assumed in more abstract models of choice behavior; indeed, it is the same equation used in the reinforcement learning model for fitting monkey's behavioral data (equation (31.2)). Our neural circuit modeling lends support to this general assumption, and allows insights into its underlying stochastic recurrent neural dynamics.

Can this model account for reaction times as well as accuracy in a decision task? This network exhibits quasi-linear ramping activity for many hundreds of milliseconds (Figure 31.9a, 31.9b). This ramping activity is realized in spite of a short neuronal membrane time constant (~10 ms) by slow reverberation mediated by the NMDA receptors (with a time constant of 50–100 ms) at the excitatory recurrent synapses. As in the drift-diffusion model, a fixed firing threshold θ can be set such that a decision is made whenever r_A or r_B reaches θ first, and the time elapsed since the stimulus onset can be read out as the decision time. This model also predicts that reaction times are longer in error trials than in correct trials, which is commonly observed in decision tasks.

Our model has been applied to a monkey experiment using a visual motion direction discrimination

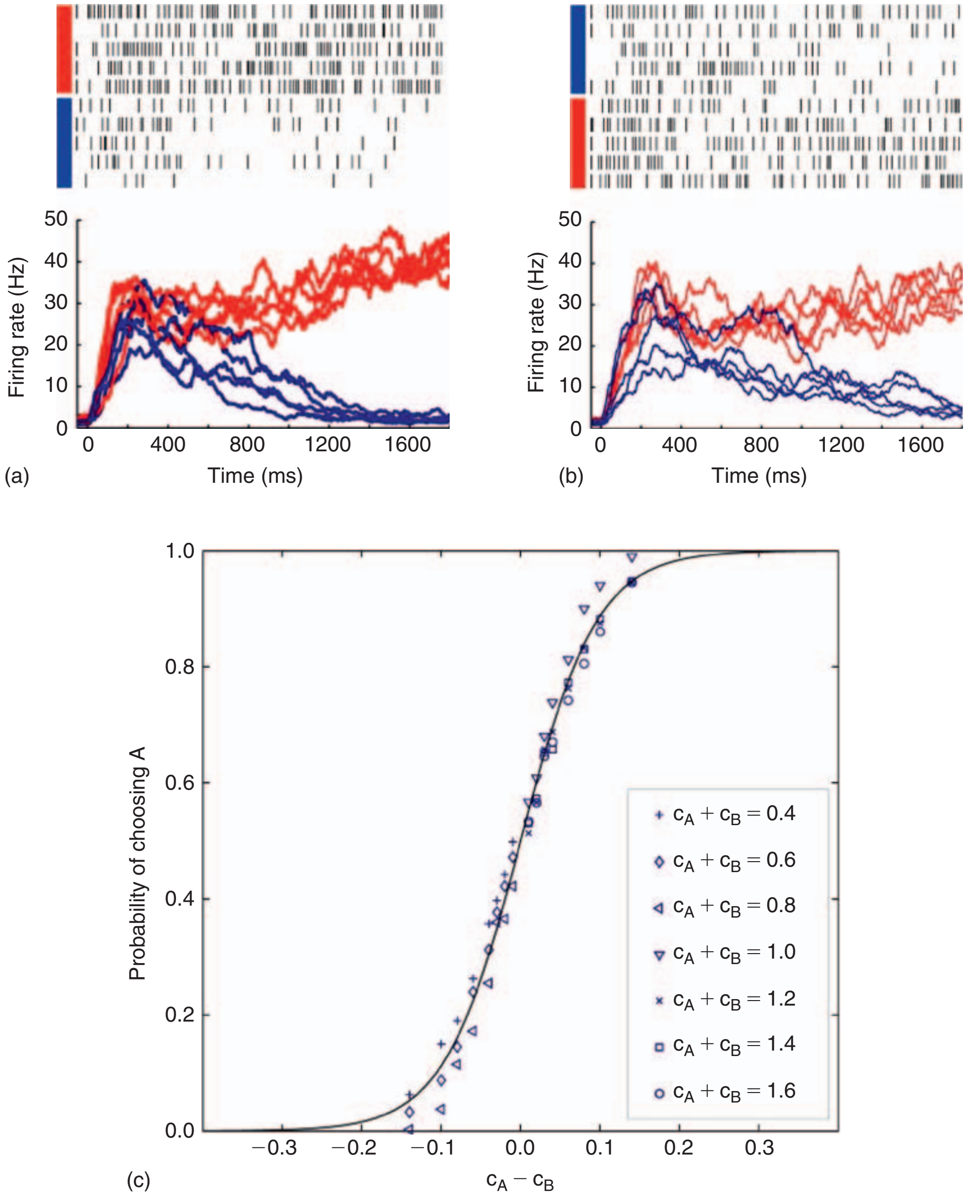


FIGURE 31.9 Neuronal activity and decision performance of the model shown in Figure 31.8. (a, b) The population firing rate of neurons is shown separately for trials in which the network's choice is the neurons' preferred (red) or non-preferred (blue) target. Raster plots show spike trains for two selected neurons in population A (top) and B (bottom) in the same trials. Neural groups A and B win the competition in trials shown in A and B, respectively. Activity is aligned at the onset of the visual targets. A few hundred milliseconds after the input onset, the average firing rates in the two populations start to diverge. Spiking activity is higher when the chosen target is preferred for the neuron (compare red and blue lines), and when its input is larger (compare red traces in the left and right panels). Moreover, firing activity is higher when the chosen target is non-preferred for the neuron that receives a larger input (compare blue traces in the left and right panels). In these simulations, the synaptic strengths are $c_A = 0.33$ and $c_B = 0.27$. (c) Choice behavior of the decision-making network for different sets of synaptic strengths. The probability of choosing target A is a function of the difference between the two synaptic strengths and it can be fitted by a sigmoid function. Different symbols represent different overall synaptic strengths.

task (Figure 31.1a). In this experiment, a subject was shown a display of moving random dots, a fraction of which moved coherently in one of two possible directions (say A = left, B = right), and the remaining dots moved in random directions. The task difficulty was varied from trial to trial by the motion coherence (0 to 100%). In monkeys performing this task, single neurons in the lateral intraparietal (LIP) cortex exhibited slow ramping activity that correlated with the perceptual decision about the direction (leftward or rightward) of the motion stimulus (Gold and Shadlen, 2007). At lower motion coherence the subject's reaction time was longer, and the ramping of the LIP neuronal firing rate was slower but reached the same firing activity level at the time when the behavioral response was produced, regardless of the motion coherence (Roitman and Shadlen, 2002). Thus, LIP neurons display a ramping-to-threshold process at the cellular level. Our neural circuit model has been used to simulate this monkey experiment, with the motion coherence given by the relative input strength $(I_A - I_B)/(I_A + I_B)$ (between -100% and $+100\%$). This model reproduces the monkey's performance and reaction times, as well as salient physiological data of LIP neurons (Wang, 2002).

Neural Substrate of a Decision Threshold

Other studies have also revealed ramping-to-threshold neural activity at the single cell level that is correlated with action selection and preparation (Schall, 2001). How can a decision threshold be instantiated by neurons, rather than prescribed in an *ad hoc* manner? A natural hypothesis is that when decision neurons integrate inputs and reach a particular firing rate level, this event triggers an all-or-none response in downstream neurons and leads to the generation of a behavioral output. This idea was tested for oculomotor decision tasks in which the motor response is a rapid saccadic eye movement. In an extended, two-stage circuit model (Lo and Wang, 2006), decision neurons in the cortex (as described above) project to movement neurons in the superior colliculus (SC), an important command center for saccades (Figure 31.10a). This model also includes a direct pathway in the basal ganglia, with an input layer (caudate, CD) and an output layer (substantia nigra reticulata, SNr). As a neural pool in the cortex ramps up in time, so do the synaptic inputs to the corresponding pool of SC movement neurons. When this input exceeds a well defined threshold level, an all-or-none burst of spikes is triggered in the movement cells, signaling a particular (A or B) motor output. In this scenario, a decision threshold (as a bound of firing rate of decision neurons) is instantiated by a hard threshold of synaptic input for downstream motor neurons. Figure 31.10b shows a sample trial of such a model simulation for the visual motion direction discrimination experiment. The rate of ramping activity fluctuates from trial to trial, as a result of stochastic firing dynamics in the cortex, and is inversely related to the decision time (as defined by the time when a burst is triggered in the SC) on a trial-by-trial basis (Figure 31.11). Moreover, when the task is more difficult (with a lower motion coherence) ramping activity is slower, leading to longer reaction times [(b) compared to (a) in Figure 31.11)]. However, the threshold of cortical

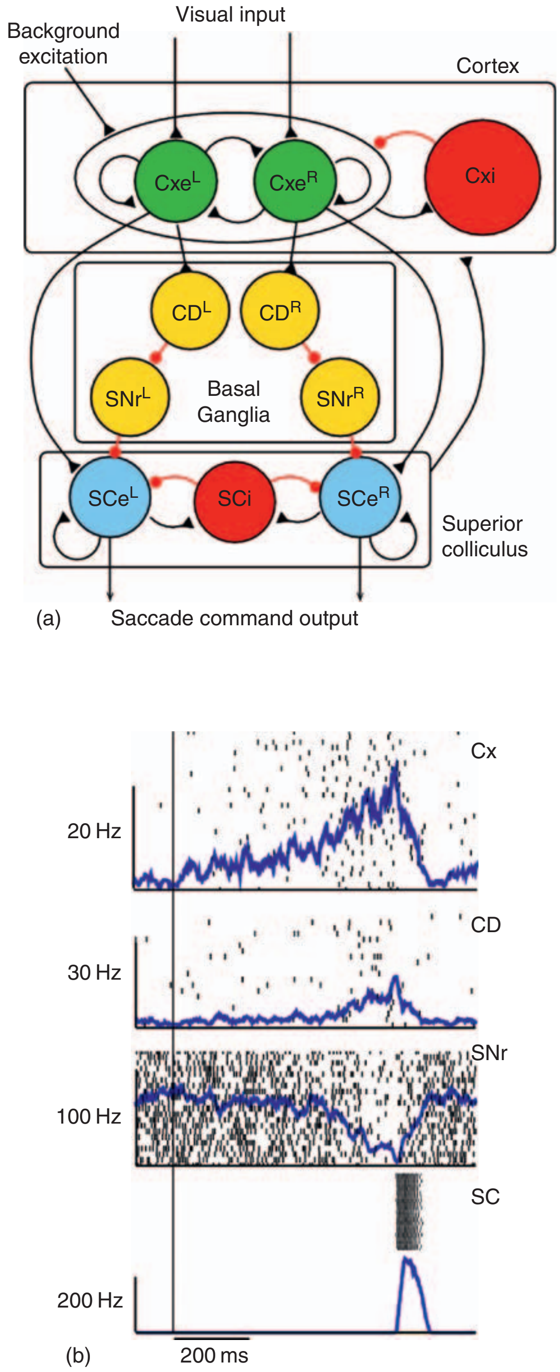


FIGURE 31.10 Multiple-circuit model of a decision threshold in the oculomotor system. (a) Schematic model architecture. Neural pools in the cortical network integrate sensory information, as well as compete against each other. They project to both superior colliculus, and caudate nucleus in basal ganglia. Excitatory and inhibitory connections are indicated as black and red lines, respectively. (b) Reaction time simulation using the oculomotor network model of spiking neurons. A saccadic response is triggered by a burst of spikes in SC movement neurons, when the ramping activity in one of the selective neural pools in the cortex exceeds a threshold, so that SC movement neurons receive sufficient cortical excitation, and are released from inhibition by the basal ganglia as a result of cortical drive to the caudate.

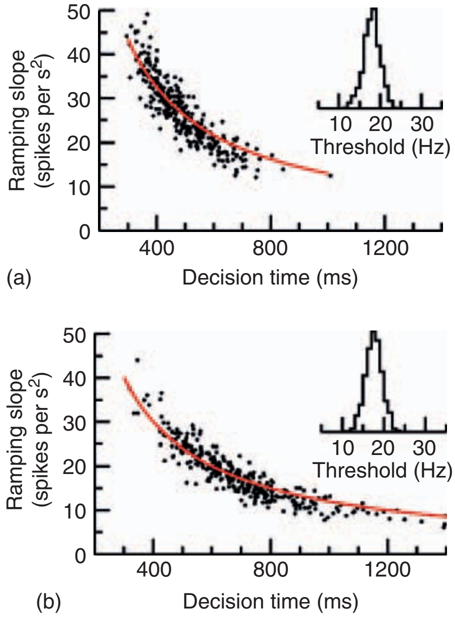


FIGURE 31.11 The slope of ramping activity of cortical decision neurons is inversely correlated with the reaction time in single trials of simulated random-dots motion direction discrimination task. When the motion strength is weaker (3.2% in (b) compared to 12.8% in (a)), the ramping slopes are smaller and the reaction times are longer. However, the variability of the cortical activity threshold remains the same (inserts), and therefore the variability in the reaction time is primarily caused by the variability of the ramping slope due to neural firing fluctuations across trials.

firing activity that is read out by the downstream motion system has the same narrow distribution (inserts in Figure 31.11), regardless of the ramping speed or reaction times. Therefore, this model realizes a robust threshold detection mechanism, and the variability of reaction times is mostly attributed to the irregular ramping of neural activity itself rather than a stochastic decision bound.

How can a decision threshold be adaptively tuned in this circuit? For instance, in a speed–accuracy tradeoff, too low a threshold leads to quicker responses but more errors, whereas too high a threshold improves the accuracy but prolongs response times. Neither of these yields maximal rewards. Since in our model the decision threshold is defined as the minimum cortical firing needed to induce a burst response in the downstream SC neurons, one would expect that this threshold could be adjusted by plastic changes in the cortico-collicular pathway: the same level of cortical input to the superior colliculus could be achieved with less firing of cortical neurons, if the synapses of the cortico-collicular projection were stronger. Interestingly, this is not the case when the system is gated by the basal ganglia. This is because neurons in SNr normally fire tonically at a high rate (Figure 31.10), and provide a sustained inhibition to SC movement neurons (Hikosaka *et al.*, 2000). This inhibition must be released, as ramping activity in the cortex activates CD neurons, which in turn suppress the activity in the SNr, in order for SC neurons to produce a burst output. This highly non-linear disinhibition mechanism implies that the decision threshold is much more readily adjustable by tuning the synaptic strength of the cortico-striatal pathway (Lo and Wang, 2006). Indeed, such an adaptive tuning of the decision threshold is expected to depend on reward signals (Reynolds *et al.*, 2001), and cortico-striatal synapses represent a major target of innervation by dopamine neurons which play a critical role in reinforcement signaling (Reynolds and Wickens, 2002). Our work suggests that dopamine-dependent plasticity of cortico-striatal synapses is a likely neural locus for adaptive tuning of the decision threshold in the brain.

REWARD-DEPENDENT PLASTICITY AND ADAPTIVE CHOICE BEHAVIOR

Computation of Returns by Synapses

To describe adaptive decision making, we have incorporated reward-dependent learning in our decision-making model. Consider a local neural network shown in Figure 31.8. Recall that the network's behavior is described by a softmax decision criterion – that is, the probability of choosing A versus B is a function of the difference in the synaptic strengths c_A and c_B for the inputs to the two competing neural pools (Figure 31.9c). Supposing that synaptic connections are plastic, then synaptic modifications will alter the network's future decision behavior, which in turn will lead to further changes in the synapses (Figure 31.8). Specifically, we used binary synapses that undergo a stochastic Hebbian learning rule, namely that synaptic plasticity depends on coactivation of presynaptic and postsynaptic neurons and takes place stochastically (Fusi, 2002). Specifically, synapses between two neurons are assumed to have two states (Down and Up). Synaptic potentiation (depression) corresponds to a probabilistic transition from the Down to Up (Up to Down) state that depends on the activity of both the pre- and post-synaptic neurons (Hebb, 1949; Fusi, 2002). In addition, it is assumed that synapses for inputs to decision neurons are potentiated only if the choice is rewarded, and depressed otherwise (Soltani and Wang, 2006; Soltani *et al.*, 2006; Fusi *et al.*, 2007). This is inspired by the suggestion that the presence or absence of dopamine signal can reverse the sign of modification at certain classes of synapses in the brain (Reynolds and Wickens, 2002). Our working hypothesis is that input synapses onto a decision circuit are updated according to such a reward-dependent Hebbian learning rule (see also Seung, 2003). As a result of synaptic modifications, the input strengths for the competing neural groups of the decision network vary from trial to trial, leading to adaptive dynamics of choice behavior.

This synaptic learning rule is a biologically plausible instantiation of reinforcement learning, and the model endowed with such synaptic plasticity is a general one rather than one designed for a particular task. We have tested the model with two types of choice behavior. First, we used the model to investigate a foraging task, in which a subject makes successive choices adaptively in a stochastic environment (Sugrue *et al.*, 2004; see also Chapter 30 of this volume). Second, the model was applied to simulate the behavior in a game-theoretic task that involves dynamic interplay between decision agents (Barraclough *et al.*, 2004), and which was summarized earlier in this chapter. In these tasks, whether a subject's choice yields reward or not depends either on the stochastic environment or on the competing agent. In either case, the model simulates a decision maker whose choice outcomes lead to synaptic plasticity that in turn influences future choices, thereby learning to forage adaptively or play a dynamic game with an opponent. We found that during learning, the synaptic strengths c_A and c_B compute returns R_A and R_B (the amount of reward per choice) rather than income (the amount of reward per unit time) for the alternative options A and B (Soltani and Wang, 2006). Moreover, because synapses are potentiated or weakened stochastically over time, they are forgetful and behave like a leaky integrator of past choice outcomes. In our model, synaptic integration of past rewards has a time constant of a few trials, and therefore the decision behavior is influenced only by rewards harvested locally in time, in agreement with behavioral (Lee *et al.*, 2004; Sugrue *et al.*, 2004; Lau and Glimcher 2005) and neurophysiological (Figure 31.6; Seo and Lee, 2007; Seo *et al.*, 2007) observations.

Matching Law: Melioration Versus Optimization

In foraging tasks commonly used in laboratories, rewards are delivered to two response options stochastically at baiting rates λ_A and λ_B, respectively, according to a particular concurrent reinforcement schedule (Figure 31.1b; Sugrue *et al.*, 2004; Lau and Glimcher, 2005). Behavioral studies using this task have led to Herrnstein's matching law, which states that a subject allocates his choices in a proportion which matches the relative reinforcement obtained from these choices (Herrnstein *et al.*, 1997). Moreover, the spiking activity of neurons in the lateral intraparietal cortex (LIP) is modulated by a representation of value that the authors defined as fractional income (Sugrue *et al.*, 2004; see also Chapter 30 for details). To explore a cortical circuit mechanism of matching behavior, we have used our neural circuit model of decision making, endowed with reward-dependent synaptic plasticity (Figure 31.8). As shown in Figure 31.12a–b, the model reproduces the matching behavior observed in the monkey experiment. As the reward rate λ_A/λ_B varies from one block of trials to the next, the choice behavior of the model changes quickly, so that the probability of choosing A versus B matches approximately λ_A/λ_B. We showed analytically that the synaptic strengths (c_A and c_B) are proportional to the returns (reward per choice) of the two targets, namely $c_A \propto R_A$ and $c_B \propto R_B$. Furthermore, the model qualitatively reproduces neural activity observed in LIP that is modulated by the values of the response options (Figure 31.13).

Figure 31.12c shows the probability of choosing option A (P_A) along with the input synaptic strengths (c_A and c_B) across six blocks of trials. The process of synaptic plasticity is stochastic, and there is considerable variability within each block of 200 trials. However, on average (indicated by the blue line for P_A), the choice probability ratio matches that of rates at which rewards are delivered to the two targets, and this matching behavior is learned through plastic synapses. For instance, if in a block of trials the reward probability λ_A is larger than λ_B, then c_A is more likely to be potentiated than c_B through the successive decisions of the network across trials because the return from choosing A is higher, leading to a larger P_A. The converse occurs in a block of trials where λ_A is smaller than λ_B.

Note that synaptic modifications take place on a trial-by-trial basis, locally in time. There is no prescription in the model for global optimization. The model's performance is close to the matching behavior, which is achieved dynamically through a melioration process – i.e., the model chooses the alternative with a higher return, so that the interplay between decision behavior and synaptic plasticity iteratively improves the total income (reward per unit time) to the maximum possible, given the constraints of the stochastic neuronal and synaptic dynamics (Soltani and Wang, 2006). Moreover, the model also reproduces the observation that, in the monkey experiment, matching is not perfect, and the relative probability of choosing the more rewarding option is slightly smaller than the relative reward rate ("under-matching") (Figure 31.12b). A model analysis showed that under-matching is a

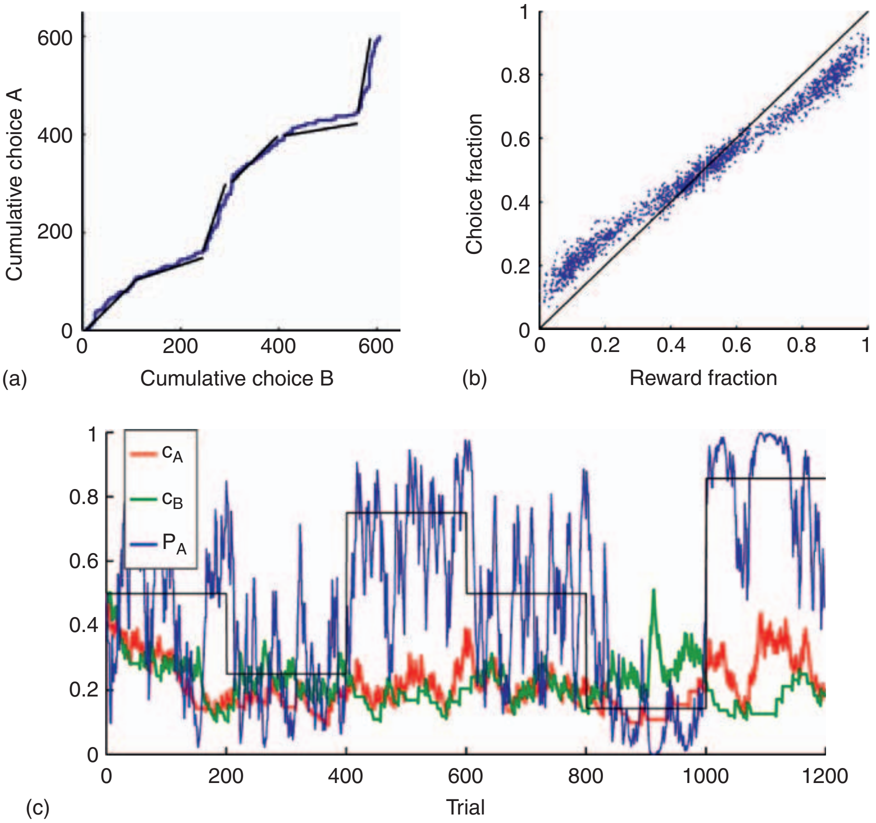


FIGURE 31.12 An attractor network model shows matching behavior in a dynamic environment. (a) For one session of the simulated matching experiment, the cumulative choice probability for target A is plotted against the cumulative choice probability for target B. Black straight lines show the baiting probability ratio in each block. The slope of the cumulative plot is equal to the choice ratio, and is approximately equal to the baiting probability ratio. In this session the following baiting probability ratios are used in sequence [1:1, 1:3, 3:1, 1:1, 1:6, 6:1]. (b) Each point shows the block-wise choice fraction as a function of the block-wise reward fraction for a block of trials on which the baiting probabilities are held constant. Most of the points fall close to the diagonal line (perfect matching), but the choice fraction is slightly lower than the reward fraction when the latter is larger than 1/2 (under-matching). (c) The synaptic strengths, c_A (red) and c_B (green), and the choice probability (blue) plotted as a function of time. The thin black line indicates the baiting probability ratio in each block. In each block the synaptic strengths fluctuate according to the value of returns from the two choices (not shown).

natural consequence of stochasticity in neural activity (Soltani and Wang, 2006).

In summary, our model suggests that although activity of LIP neurons depends on values of response options (Figure 31.13), valuation may occur elsewhere, at the synaptic level and in the form of returns. For the sake of simplicity we have considered a local network model, but we are agnostic regarding the actual site of synaptic plasticity that is critically involved with valuation. Candidate loci include the cortico-striatal connections in the basal ganglia (Lo and Wang, 2006), or synaptic pathways within the orbitofrontal cortex (see Chapter 29).

Random Choice Behavior in Matching Pennies Game

Our model can be extended to decision making in competitive games between multiple agents, which is the main topic of the earlier sections of this chapter. The idea is that several such models, each simulating a "decision maker," can interact according to a payoff matrix. We have used our model to simulate the monkey experiment of Barraclough *et al*. (2004), in which monkeys play matching pennies with a computer opponent that uses three different algorithms (0, 1, and 2, see description above). The model reproduces many salient behavioral observations (Soltani *et al*., 2006). If the opponent is not interactive (using Algorithm 0), the model decision behavior is idiosyncratic and might, for instance, choose one of the targets exclusively. When the opponent uses algorithm 1, the model exhibits prominent win–stay–lose–switch (WSLS) behavior, as observed in monkeys. Finally, when the opponent uses algorithm 2 and is fully interactive according to the rules of matching pennies, the model behavior becomes quasi-random. This is shown in Figure 31.14, with several different sets of initial

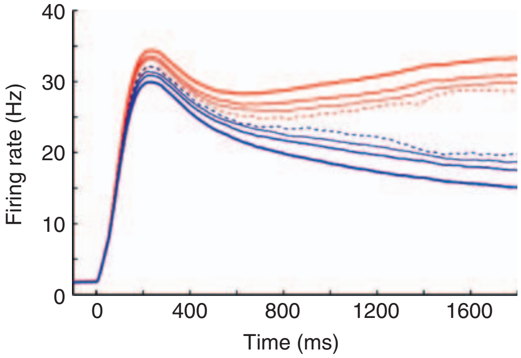
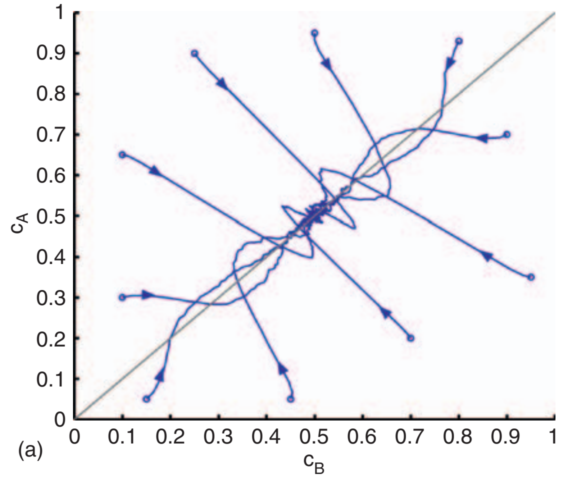


FIGURE 31.13 Graded activity of model neurons as a function of the input synaptic strengths which encode the values (returns) of choice options in a matching task. The activity of decision neurons shows a graded pattern, if single-trial firing rates are sorted and averaged according to the network's choice and the difference between synaptic strengths. Activity is aligned by the onset of two targets, and is shown separately for the choices corresponding to the neurons' preferred (red) or non-preferred (blue) target. In addition, trials are subdivided into four groups according to the difference between the strength of synapses to the two competing neural populations [$c_A - c_B = -0.16$ to -0.05 (dashed), -0.05 to 0 (thin), 0 to 0.05 (medium), 0.05 to 0.16 (thick)].

values for the synaptic variables c_A and c_B (Figure 31.14, (a)). Different c_A and c_B values yield a different initial probability P_A of choosing response A versus B (Figure 31.14, (b)). Competitive interaction with the opponent, however, quickly equalizes the synaptic variables (Figure 31.14, (a)), and the choice probability becomes very close to 0.5 (Figure 31.14, (b)), regardless of the initial state of the system. For instance, if initially the system chooses target A more frequently because c_A is larger than c_B, it would be exploited by the opponent, and the unrewarded outcomes from choosing A would induce depression of c_A of the synapses to the neural pool A, so that the difference $c_A - c_B$ decreases over time, and the system gradually chooses B more frequently.

Interestingly, our model, with a reinforcement learning rule that changes only synapses onto the neurons selective for the chosen option, does not capture all the details of the monkey's behavior. In particular, it shows a probability of WSLS, P(WSLS), below a limited value (about 0.65), whereas P(WLSL) can be nearly 1 in monkeys with algorithm 1. We found that P(WLSL) ~1 can be realized in our model with a different learning rule, according to which synapses onto both neural populations (selective for the chosen and unchosen targets) are modified in each trial. This is akin to a "belief-dependent learning rule" (Camerer, 2003; Lee *et al.*, 2005). It is also in line with the conclusion that both past reinforcers and past choices influence future decision behavior (Lau and Glimcher, 2005).



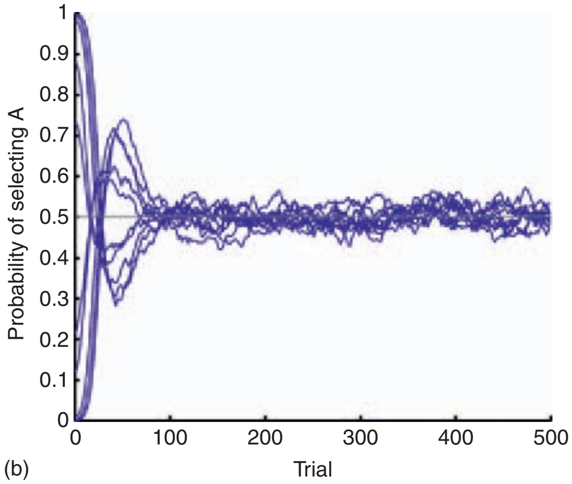


FIGURE 31.14 Model simulation of the dynamic choice behavior during the matching pennies game. (a) The synaptic strengths c_A and c_B plotted against each other show the time evolution of adaptive learning. (b) The corresponding choice probability for A, which is a softmax function of $c_A - c_B$. Regardless of the initial state of the neural circuit (different c_A and c_B values), the network quickly evolves towards the Nash Equilibrium state of random mixed strategy, with $c_A \approx c_B$ (the diagonal line in (a)), and the choice probability becomes chance level (0.5 in (b)).

Although our model can reproduce monkey's behavior obtained with different algorithms, different model parameters are required for each algorithm. How can these model parameters be tuned adaptively, as the opponent's algorithm is changed? To address this question, we incorporated a meta-learning rule proposed by Schweighofer and Doya (2003) that maximizes long-term rewards. We found that the enhanced model captures the very slow changes of the monkey's behavior, as the opponent's algorithm changes from session to session (Soltani *et al.*, 2006).

A general insight of this work is that a decision circuit produces random choice behavior, not necessarily because the system has a prescribed "random number generator," but because the trial-to-trial choice dynamics force the decision agent to play randomly. This is well demonstrated in our model, because the same model produces either stereotypical responses or random responses, depending on the behavior of its opponent. The model decision maker does not have the goal to play randomly, but simply tries to play its best, given the environment and other decision agents involved in the game. This conclusion is consistent with previous behavioral studies and models, emphasizing the critical importance of feedbacks in the production of quasi-random behavior (Rapoport and Budescu, 1992; Camerer, 2003). Moreover, our model suggests that irregular neural firing that gives rise to sigmoid decision criterion, and the stochastic nature of synaptic learning, contribute to the generation of random choice behavior, which can be desirable and even optimal in interactive decision tasks. Thus, this model sheds insights into neural processes in the brain underlying the randomness observed at the psychological level (Glimcher, 2005). Therefore, neurobiologically based neural modeling helps to bridge the gap between cognitive behavior and its underlying neural network mechanisms.

CONCLUSION

Much of the research in behavioral economics focuses on how decision makers choose among various options when the information about the uncertain future prospects are provided explicitly. For example, in studies on decision making under risk, the decision makers are given specific information about the magnitudes and probabilities of possible payoffs from each choice. In real life, however, information about the magnitude, likelihood, and temporal delay of reward and punishment resulting from a particular choice often has to be estimated through experience. Furthermore, such reward contingencies often change over time, and this happens frequently when multiple agents interact. Especially valuable insight into the cognitive processes underlying such adaptive choice behaviors comes from studies on choice behaviors in games. Simple competitive zero-sum games, such as matching pennies and Rock–Paper–Scissors, also provide opportunities to study the underlying neural mechanisms for dynamic decision making, because non-human primates can be easily trained to play such games against a computer-simulated opponent.

Recently, neurophysiological studies have identified different types of signals encoded by individual neurons in the frontal cortex and the posterior parietal cortex during such computer-simulated competitive games that might be used to determine the choice of the animal and update its strategy based on the previous choice outcomes. These studies have largely focused on the neural correlates of value functions and the signals necessary to update the value functions according to simple reinforcement learning algorithms. In contrast, the behavioral strategies of humans and animals might be also influenced by high-level cognitive factors, such as abstract rules and behavioral contexts, and the knowledge of the dynamic properties of the environment (Hampton *et al.*, 2006). The neural mechanisms responsible for updating the animal's decision-making strategies based on such multiple sources of information are not well understood, and should be pursued in future research. Future studies also need to elucidate the neural processes involved in integrating different types of costs and benefits during decision making (Roesch *et al.*, 2006; Lee *et al.*, 2007; Rushworth *et al.*, 2007). It might be also possible to investigate the neural processes specialized for social interactions, since non-human primates can be also trained to play relatively simple experimental games in laboratory settings (Chen and Hauser, 2005; Jensen *et al.*, 2007).

The results from neurophysiological recording experiments need to be linked to mechanistic models about how information regarding the outcomes of previous choices is incorporated into a network of spiking neurons, allowing the animal to adjust its decision-making strategies adaptively. We have developed such a model of spiking neurons that is characterized by strongly recurrent or attractor dynamics and endowed with reward-dependent Hebbian synaptic plasticity. These studies have begun to provide important clues as to how adaptive stochastic decision making, such as matching behavior in a foraging task or approximate Nash Equilibrium in a competitive game, result from a dynamic interplay between a decision-making network and its environment. This model will need to be extended to investigate how a neural network or system of networks can optimally combine the information about various aspects of reward and punishment, such as their magnitude, probability, and immediacy. Also, the biophysical basis of reward-dependent plasticity in the brain remains to be elucidated. We expect that progress in this direction will ultimately account for the discrepancy between the choice behaviors of humans and animals, and the rational choice behaviors prescribed by normative economic theories.

Acknowledgments

We thank Dominic Barraclough and Hyojung Seo for their contributions to the neurophysiologial studies, Chung-Chuan Lo and Alireza Soltani for their contributions to the computational modeling and help with the figure and Alireza Soltani for his comments on the manuscript. This research was supported by NIH grants MH073246 (D.L. & X.W.) and NS044270 (D.L.).

References

Amemori, K. and Sawaguchi, T. (2006). Contrasting effects of reward expectation on sensory and motor memories in primate prefrontal neurons. *Cerebral Cortex* 16, 1002–1015.

Amiez, C., Joseph, J.P., and Procyk, E. (2005). Anterior cingulate error-related activity is modulated by predicted reward. *Eur. J. Neurosci.* 21, 3447–3452.

Amiez, C., Joseph, J.P., and Procyk, E. (2006). Reward encoding in the monkey anterior cingulate cortex. *Cerebral Cortex* 16, 1040–1055.

Amit, D.J. and Mongillo, G. (2003). Selective delay activity in the cortex: phenomena and interpretation. *Cerebral Cortex* 13, 1139–1150.

Barbas, H. (1995). Anatomical basis of cognitive-emotional interactions in the primate prefrontal cortex. *Neurosci. Biobehav. Rev.* 19, 499–510.

Barbas, H. and De Olmos, J. (1990). Projections from the amygdala to basventral and mediodorsal prefrontal regions in the rhesus monkey. *J. Comp. Neurol.* 300, 549–571.

Barbas, H. and Pandya, D.N. (1987). Architecture and frontal cortical connections of the premotor cortex (area 6) in the rhesus monkey. *J. Comp. Neurol.* 256, 211–228.

Barbas, H. and Pandya, D.N. (1989). Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.* 286, 353–375.

Barraclough, D.J., Conroy, M.L., and Lee, D. (2004). Prefrontal cortex and decision-making in a mixed-strategy game. *Nat. Neurosci.* 7, 404–410.

Bogacz, R., Brown, E., Moehlis, J. *et al*. (2006). The physics of optimal decision-making: a formal analysis of models of performance in two-alternative forced choice tasks. *Psychol. Rev.* 113, 700–765.

Camerer, C.F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Carmichael, S.T. and Price, J.L. (1996). Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *J. Comp. Neurol.* 371, 179–207.

Chafee, M.V. and Goldman-Rakic, P.S. (1998). Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task. *J. Neurophysiol.* 79, 2919–2940.

Chen, M.K. and Hauser, M. (2005). Modeling reciprocation and cooperation in primates: evidence for a punishing strategy. *J. Theoretical Biol.* 235, 5–12.

Constantinidis, C., Franowicz, M.N., and Goldman-Rakic, P.S. (2001). The sensory nature of mnemonic representation in the primate prefrontal cortex. *Nat. Neurosci.* 4, 311–316.

Dorris, M.C. and Glimcher, P.W. (2004). Activity in the posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44, 365–378.

Douglas, R.J. and Martin, K.A.C. (2004). Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* 27, 419–451.

Fusi, S. (2002). Hebbian spike-driven synaptic plasticity for learning patterns of mean firing rates. *Biol. Cybern.* 87, 459–470.

Fusi, S., Asaad, W., Miller, E.K., and Wang, X.-J. (2007). A neural model of flexible sensori-motor mapping: learning and forgetting on multiple timescales. *Neuron* 54, 319–333.

Fuster, J.M. (1997). *The Prefrontal Cortex: Anatomy, Physiology and Neuropsychology of the Frontal Lobe*. Philadelphia, PA: Lippincott-Raven Publishers.

Ghashghaei, H.T., Hilgetag, C.C., and Barbas, H. (2007). Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala. *NeuroImage* 34, 905–923.

Glimcher, P.W. (2005). Indeterminacy in brain and behavior. *Annu. Rev. Psych.* 56, 25–56.

Gold, J.I. and Shadlen, M.N. (2007). The neural basis of decision-making. *Annu. Rev. Neurosci.* 30, 535–574.

Hampton, A.N., Bossaerts, P., and O'Doherty, J.P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision-making in humans. *J. Neurosci.* 26, 8360–8367.

Hastie, H., Tibshirani, R., and Friedman, J. (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer.

Hebb, D.O. (1949). *Organization of Behavior*. New York, NY: Wiley.

Herrnstein, R.J., Rachlin, H., and Laibson, D.I. (1997). *The Matching Law: Papers in Psychology and Economics*. Cambridge, MA: Harvard University Press.

Hikosaka, O., Takikawa, Y., and Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol. Rev.* 80, 953–978.

Hoshi, E., Sawamura, H., and Tanji, J. (2005). Neurons in the rostral cingulate motor area monitor multiple phases of visuomotor behavior with modest parametric selectivity. *J. Neurophysiol.* 94, 640–656.

Jensen, K., Call, J., and Tomasello, M. (2007). Chimpanzees are rational maximizers in an ultimatum game. *Science* 318, 107–109.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–292.

Kandel, E.R., Schwartz, J.H., and Jessel, T.M. (2000). *Principles of Neural Science*. New York, NY: McGraw-Hill.

Kobayashi, S., Lauwereyns, J., Koizumi, M. *et al*. (2002). Influence of reward expectation on visuomotor processing macaque lateral prefrontal cortex. *J. Neurophysiol.* 87, 1488–1498.

Lau, B. and Glimcher, P.W. (2005). Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* 84, 555–579.

Lee, D., Conroy, M.L., McGreevy, B.P., and Barraclough, D.J. (2004). Reinforcement learning and decision-making in monkeys during a competitive game. *Cogn. Brain Res.* 22, 45–58.

Lee, D., McGreevy, B.P., and Barraclough, D.J. (2005). Learning and decision-making in monkeys during a rock–paper–scissors game. *Cogn. Brain Res.* 25, 416–430.

Lee, D., Rushworth, M.F.S., Walton, M.E. *et al*. (2007). Functional specialization of the primate frontal cortex during decision-making. *J. Neurosci.* 27, 8170–8173.

Leon, M.I. and Shadlen, M.N. (1999). Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* 24, 415–425.

Lewis, D.A., Melchitzky, D.S., Sesack, S.R. *et al*. (2001). Dopamine transporter immunoreactivity in monkey cerebral cortex: regional, laminar, and ultrastructural localization. *J. Comp. Neurol.* 432, 119–136.

Lo, C.C. and Wang, X.-J. (2006). Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat. Neurosci.* 9, 956–963.

Luppino, G., Rozzi, S., Calzavara, R., and Matelli, M. (2003). Prefrontal and agranular cingulate projections to the dorsal premotor areas F2 and F7 in the macaque monkey, Eur. *J. Neurosci.* 17, 559–578.

Machens, C.K., Romo, R., and Brody, C.D. (2005). Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science* 18, 1121–1124.

Matsumoto, M., Matsumoto, K., Abe, H., and Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nat. Neurosci.* 10, 647–656.

Miller, P. and Wang, X.-J. (2006). Inhibitory control by an integral feedback signal in prefrontal cortex: a model of discrimination between sequential stimuli. *Proc. Natl Acad. Sci. USA,* 103, 201–206.

Nash, J.F., Jr. (1950). Equilibrium points in n–person games. *Proc. Natl Acad. Sci. USA* 36, 48–49.

Niki, H. and Watanabe, M. (1979). Prefrontal and cingulate unit-activity during timing behavior in the monkey. *Brain Res.* 171, 213–224.

Pawitan, Y. (2001). *In All Likelihood: Statistical Modelling and Inference using Likelihood*. Oxford: Oxford University Press.

Rapoport, A. and Budescu, D.V. (1992). Generation of random series in two-person strictly competitive games. *J. Exp. Psych. Gen.* 121, 352–363.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.* 85, 59–108.

Reynolds, J.N. and Wickens, J.R. (2002). Dopamine-dependent plasticity of corticostriatal synapses. *Neural Networks* 15, 507–521.

Reynolds, J.N., Hyland, B.I., and Wickens, J.R. (2001). A cellular mechanism of reward-related learning. *Nature* 413, 67–70.

Roesch, M.R., Taylor, A.R., and Schoenbaum, G. (2006). Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* 51, 509–520.

Roitman, J.D. and Shadlen, M.N. (2002). Responses of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489.

Rushworth, M.F.S., Behrens, T.E.J., Rodebeck, P.H., and Walton, M.E. (2007). Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends Cogn. Sci.* 11, 168–176.

Schall, J.D. (2001). Neural basis of deciding, choosing and acting. *Nat. Rev. Neurosci.* 2, 33–42.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.

Schweighofer, N. and Doya, K. (2003). Meta-learning in reinforcement learning. *Neural Networks* 16, 5–9.

Seo, H. and Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* 27, 8366–8377.

Seo, H., Barraclough, D.J., and Lee, D. (2007). Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex. *Cerebral Cortex* 17, i110–117.

Seung, H.S. (2003). Learning in spiking neural networks by reinforcement of stochastic synaptic transmission. *Neuron* 40, 1063–1073.

Smith, P.L. and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161–168.

Soltani, A. and Wang, X.-J. (2006). A biophysically based neural model of matching law behavior: melioration by stochastic synapses. *J. Neurosci.* 26, 3731–3744.

Soltani, A., Lee, D., and Wang, X.-J. (2006). Neural mechanism for stochastic behavior during a competitive game. *Neural Networks* 19, 1075–1090.

Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.

Sutton, R.S. and Barto, A.G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Usher, M. and McClelland, J.L. (2001). On the time course of perceptual choice: the leaky competing accumulator model. *Psychol. Rev.* 108, 550–592.

Vogt, B.A. and Pandya, D.N. (1987). Cingulate cortex of the rhesus monkey: II. Cortical afferents. *J. Comp. Neurol.* 262, 271–289.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Wald, A. and Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. *Ann. Math. Stats,* 19, 326–337.

Wang, X.–J. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.* 24, 455–463.

Wang, X.-J. (2002). Probabilistic decision-making by slow reverberation in cortical circuits. *Neuron* 36, 955–968.

Wang, Y., Shima, K., Sawamura, H., and Tanji, J. (2001). Spatial distribution of cingulate cells projecting to the primary, supplementary, and pre-supplementary motor areas: a retrograde multiple labeling study in the macaque monkey. *Neurosci. Res.* 39, 39–49.

Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature* 382, 629–632.

Williams, S.M. and Goldman-Rakic, P.S. (1993). Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. *Cerebral Cortex* 3, 199–222.

Wong, K.-F. and Wang, X.-J. (2006). A recurrent network mechanism for time integration in perceptual decisions. *J. Neurosci.* 26, 1314–1328.

CHAPTER

32

Choice: Towards a Standard Back-pocket Model

Paul W. Glimcher

OUTLINE

INTRODUCTION

The goal of neuroeconomics is an algorithmic description of the human mechanism for choice. How far have we proceeded towards this goal? This volume reveals just how much information has been gathered. The studies presented here have leveraged existing scholarship to describe the mechanisms by which the values of actions are learned, how and where these values are encoded, how these valuations govern our actions, and how neural measurements can be used to constrain social scientific models of human behavior. With this information in hand, can we define the gross features of the human choice system in a way that will be of use to economists, psychologists, and neuroscientists? Or, to put it more precisely, can we use the lens of economic theory and experiment to better understand the neurobiological and psychological data in a way that will benefit all three disciplines? My suspicion is that many of the scholars contributing to this volume would say that the answer to this question is yes, and that the back-pocket models that most of these scholars use to guide their research are remarkably similar.

With that in mind, this chapter seeks a fairly precise definition of a standard back-pocket model of human decision making that incorporates the bulk of what we know today. Of course such a model, if taken seriously by policy-makers, could be dangerous. Many of the details of such a framework, even if all of those details were supported by contemporary data, would be both controversial and wrong. But a framework that organizes the information we have, and serves as a target for future challenges, may well maximize the forward movement of our discipline. Such a scaffold might make clearer what we do know and what we do not know; where we have made progress, and where critical avenues remain unexplored. In that spirit, and with the certain knowledge that the details of the following framework are wrong, what follows is a fairly formal presentation of a "standard back-pocket model" for choice.

My goal in presenting this model is to explicitly link neurobiological, psychological, and economic studies of choice so that we can examine the implications of this structure for all three of our parent disciplines. Simply relating a chooser's options to her choices or simply specifying patterns of brain connectivity and activation would run counter to the goals and spirit of neuroeconomics: What is called for is a hybrid approach that rigorously mixes the strategies and traditions of our fields so as to explicitly maximize the number of constraints these parent disciplines can impose on our understanding of choice. In interpreting this approach, scholars trained in only one of the parent disciplines may be initially troubled. For classical neurobiologists, this approach may seem to include an overly formal definition of conceptual objects to no particular end. For economists, the emphasis may seem overly algorithmic and unnecessarily focused on cardinality. My own feeling, however, is that the interaction of these constraints is what makes neuroeconomics powerful. As I hope will become clear in this chapter, the explicit ties to economics will allow neurobiologists to rule out whole classes of theories that have heretofore seemed reasonable. The explicit ties to neurobiological data will reveal that only a tiny space in the vast landscape of economic theory can be viewed as compatible with the human neuroarchitecture. In any case, I ask the reader's forbearance in this regard. I mean the presentation to be *neuroeconomic*. I hope the presentation will reveal two things: (1) that only a very specific set of economic theories is supported by the available data, and (2) that several very specific pieces of neurobiological and psychological data are required to complete the theoretical constraints on the architecture of human choice.

What follows, then, is a presentation in five parts. The first section provides a quick overview of the basic mechanism for which I intend to argue. This section is not meant to be a defensible piece of evidence-based reasoning, but rather an opportunity to sketch out the shape of the coming argument. The second section provides a formal definition of the mathematical and empirical objects used in the rest of the presentation. For some this may seem superfluous, and for others it may seem *ad hoc* or overly restrictive. I hope it will become clear as the exposition develops that we require these particular objects to link existing economic theory to empirical psychological and neuroscientific data. The third section will provide a detailed description of the evidence for a generalized neural mechanism of valuation – a detailed description of what we do and do not know about this process, and a description of its surprisingly unitary and linear nature. The fourth section provides an overview of the choice mechanism itself; those circuits that take as their inputs the outputs of the valuation system and give as their output a plan of action – a choice. The chapter concludes by highlighting both the strengths and weaknesses of this standard back-pocket model.

THE BASIC TWO-STAGE MODEL

Growing evidence suggests that the basic mechanism for producing choices in primates of all kinds (a group which necessarily includes humans) involves a two-stage mechanism. The first of these stages is concerned with the valuation of all goods and actions; the second is concerned with choosing amongst the goods or actions presented in a given choice set. At a very basic level, one can think of the valuation mechanism as being associated with learning and representing the values of objects and actions. The choice mechanism can be viewed performing a transformation that takes as an input the values of the options under current consideration – the choice set – and stochastically returns a high-valued option used to guide physical action. Of course, the details of these mechanisms are subtle. Some features of the valuations we infer from behavior (what an economist would call the *preference function*) seem to be attributable to mechanical processes embedded in the choice mechanism itself. For example, the model suggests that hyperbolic temporal discounting may arise from a mixture of exponential temporal discounting within the valuation system and a set of divisive computations embedded in the choice mechanism (see, for example, Glimcher *et al.*, 2007). The mapping between the physical valuation mechanisms of the brain and psychological notions of valuation will occasionally be complicated. In a similar way, the anatomical boundaries between the choice and valuation mechanisms may not be entirely discrete. Our mixture of theoretical and empirical approaches will make it clear that neural activity in the *choice structures* both should and does influence activity in the valuation structures. For example, some neurons in key valuation areas like the striatum carry signals that encode choice. However, at a global level, it now seems extremely likely that the architecture is organized around this basic two-stage framework.

The neurobiological evidence for a two-stage process, which will be reviewed in greater detail below, arises from several key observations which are summarized only very briefly here. Perhaps the first explicit evidence for this segregation came from the work of Platt and Gimcher (1999). In the first half of that study, the authors recorded from neurons in the posterior parietal cortex while thirsty monkeys viewed two visual targets. In a typical trial, or round, the two targets might be associated with different magnitudes of reward, and after an initial delay the animal was informed which one of the two targets would yield that reward on this trial. From a choice-related point of view, each round was a decision between a response that yielded no reward and a response that yielded a small positively-valued reward. What Platt and Glimcher found, however, was that during the early part of the trial (before the zero-valued target was identified for that trial) these neurons produced firing rates almost linearly proportional to the average value of the rewards that had previously been earned for selecting that target. These firing rates cardinally encoded, in action potentials (or spikes) per second, the average value of the targets, but in a way that *did not* influence choice on that trial. It was "as if" the mean expected utility of the action *look at the right target* was linearly encoded by neuronal firing rates independent of choice. This is a point that will be developed in greater detail below, but the point I want to make here is that at the time this was seen as a major limitation of the study. In retrospect it provides some of the first compelling evidence that valuation and choice are dissociable.

At the same time, a huge number of studies from many different sub-areas of neuroscience began to suggest that broad swaths of the striatum and the frontal cortex both learn and represent the values of goods and actions even when learning is passive (Figure 32.1). Delgado *et al.* (2000) and Knutson *et al.* (2000), for example, found that when humans passively viewed events that resulted in unpredictable gains or losses, the level of neural activation in several striatal and fronto-cortical areas was linearly correlated with the magnitudes of these gains and losses. This constituted preliminary evidence that these value-encoding structures operate in the absence of choice[1]. Subsequent studies confirm this initial observation. During both choice and non-choice tasks, when humans face risky or certain gains, when they face delayed or immediate gains, the activation of discrete frontal and striatal nuclei is almost always near-linearly related to (subjective) measures of value. The linearity of this relationship may be surprising to economists, but this is not theory, it is simply an empirical (and surprisingly common neuroscientific) observation. These existing studies have made it clear that the neural systems for valuation are both neurochemically and anatomically localized.

The critical first step towards this realization was the identification of reinforcement learning mechanisms in the forebrain, and it is an understanding of these learning mechanisms that has paved the way towards a broader understanding of valuation. In the early 1990s, Wolfram Schultz and his colleagues (see, for example, Romo and Schultz, 1990; Schultz and Romo, 1990; Schultz, *et al*., 1993; also Chapter 21 in this volume) demonstrated that midbrain dopaminergic neurons encode a *reward-prediction error*. These are highly specialized and anatomically localized neurons that broadcast a signal throughout the striatum and the frontal cortex. Montague and colleagues (1997; see also Chapter 22 in this volume) provided the next step when they recognized that this class of signal could be used to construct a mechanism that learns, through trial-and-error, the values of actions or objects. What followed was 10 years of work which established the existence of at least three inter-related subsystems in these brain areas that employ distinct mechanisms for learning and representing value, and which interact to produce the valuations that guide choice (as summarized in Chapters 12 and 24). This provided a set of landmark findings summarized by the many chapters in Part 3 of this volume.

In a similar way, studies of the movement control systems of the brain revealed both the need for and existence of a discrete choice mechanism. Although it may not be obvious to non-neuroscientists, the process of producing a movement at the biomechanical level is extremely complicated. Checking one of two boxes on a retirement fund contract, or signaling

---

[1] Of course, this analysis presumes that experiencing rewards and anticipating them for the purposes of decision share a common neural substrate. That this is true is now largely beyond dispute at the neurobiological level – a point that is developed later in the chapter.

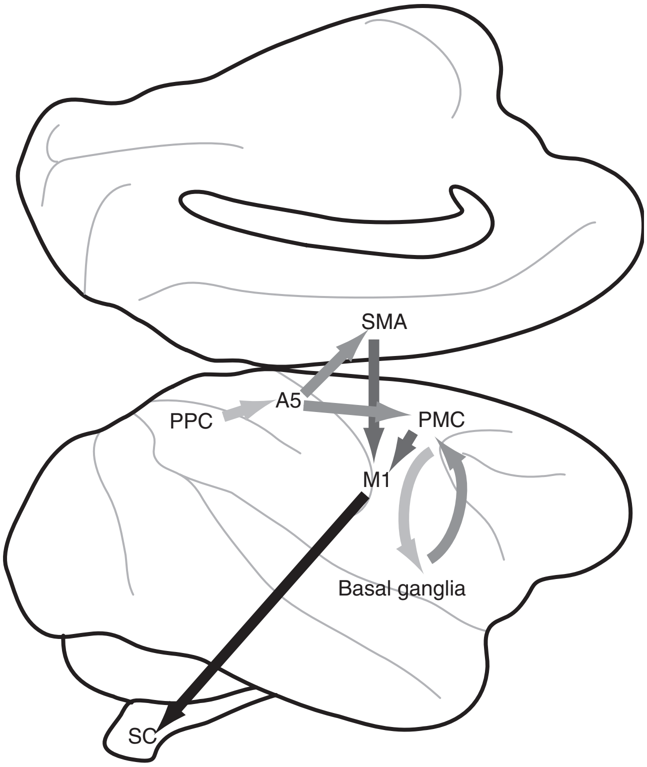


FIGURE 32.1 A highly simplified view of the skeleto-motor output system of the monkey brain. The key feature communicated by the figure is that behavioral outputs must, in principle, converge on a final common pathway for movement control. PPC, posterior parietal cortex; A5, Broadmann's Area 5; SMA, supplementary motor area; PMC, premotor cortex; M1, motor cortex.

a choice with a rightward eye movement, requires quite precise coordination of what may be literally dozens of independent muscles. Each of these movements, though, is a unitary object that must be selected and planned before it is executed. The hypothesis that option values influence muscles directly is easily falsified. When a subject moves her pencil towards a checkbox on a page, the tip of the pencil is moved by over 30 muscles and joints with more than 7 degrees of freedom. Still, the tip of that pencil traces a straight line from start to checkbox with a Gaussian velocity profile that peaks about halfway through the movement. Reaching for that checkbox reflects a movement richly planned before it is executed. The most introductory neuroscience textbook reveals this fact. Of course, we know where much of the neural hardware that plans and regulates movements is located. Areas like the motor cortex, the premotor cortex, the supplementary motor area, and Broadmann's area 5 all coordinate the generation of goal directed movements of the arms. A similar (and better understood) system coordinates movements of the eyes. What this means is that value signals must be turned into action control signals somewhere within the nervous system, presumably at or above the level of motor cortex. It is this process that neurobiologists refer to as a *choice mechanism*. What I point out here is that it is critical to keep in mind that choice must be accomplished before movements are designed. The movement control systems reflect a *final common path* before which choice must, under normal circumstances, be complete.

Our current evidence suggests that the choice system involves large portions of the parietal cortex, amongst other areas (as summarized in Part 5 of this volume). These parietal areas receive both direct and indirect projections from the valuation areas, and project directly to the movement control areas. One issue that remains unclear, however, is how much the frontal cortex and basal ganglia participate directly in the choice process working in concert with these parietal areas. We now know that specific neurons in the orbitofrontal cortex (as reviewed here in Chapter 29) and the dorsal striatum (Samejima, *et al.*, 2005; Lau and Glimcher, 2008) of the monkey also represent goods and actions that have been chosen before these choices are executed, but whether these neurons participate directly in choice is not known at this time.

This then, is a minimal working outline of the primate choice system: a valuation system that learns through repeated sampling of the environment and stores the values of actions and/or goods; a choice system that uses these values to select (from amongst a current choice set) a single option; and a motor control system that executes the physical responses dictated by the choice. Of course future experiments will enrich this description – for example, it may well be the case that perceptual systems influence the valuation systems in ways that we are just beginning to understand – but these seem to be the fundamental components of the primate architecture for choice as we understand it today.

Before beginning to examine the valuation and choice systems in detail, however, it is critical that we link these components to economic theory. This will provide both important constraints on how these systems operate and a common language for thinking about these problems. In what follows, definitions for conceptual objects explicitly linked to economic theory are presented. While these definitions may initially seem opaque to neurobiologists, neurobiological readers are urged to take them seriously. The role of these objects is to serve, in essence, as mapping rules that connect existing theoretical tools to the empirical measurements of neuroscience. At an earlier point in the history of neuroeconomics, it may have been valuable to speak in broadly metaphorical terms when saying things like "this neuronal firing rate is like a utility signal[2]". But as this field transitions to formal tests of explicit and powerful theories, this kind of metaphorical relationship between theory and data becomes more and more untenable. If we are to leverage the precise and highly testable (if rarely tested) theories of economics and psychology, then we must be able to specify clearly *how* we would test those theories with neurobiological data – not just in a general way, but in a specific computational sense. The kinds of objects described below, I argue here, are what is required.

To many economists, a feature that will stand out about these objects is that they will seem unnecessarily restrictive. Linear relationships will be postulated which do not seem necessary. From a logical point of view, I agree. These, however, are the objects for which our empirical data calls. To many neurobiologists, the objects I have selected for definition may seem arbitrary. In the sections below, I hope to convince you that the empirical data argue that these are also called for by our current data.

DEFINING OBJECTS

Expected Utility Theory

As several chapters in this volume make clear, the axiomatic approach in general and expected utility theory in particular have both good and bad features (for an overview of its advantages, see Chapter 3; for an overview of its weaknesses see Chapter 11). Formally, the theory of expected utility (von Neumann and Morgenstern, 1944) rests on four axioms (or three, in Savage's 1954 formulation). For our purposes, I want to stress why these axioms are not some set of strange and arbitrary assumptions about how people *must* behave, which is an interpretation often given to them by critics. The axioms are a statement not about people (or the brain) in any sense; the axioms are a precise definition of a theory. It is reasonable to dislike any theory, but it is important to stress that, counter to what many lay people believe, this is a very minimalistic theory – much, much less restrictive and much more intuitive than, for example, temporal difference theories of learning. If what neurobiologists studying decision making want is a simple theory of how people value things, then it is important for them to realize that economists already have several such theories, and that the implications of these theories have been very well explored. When a class of behavior obeys the axioms of a given economic theory, then we already know quite a bit about the valuation systems that can, in principle, underlie that behavior. It is for this reason that neurobiologists need to link their measurements to economic frameworks. This is the only way that neurobiologists can rigorously exploit what economists have already learned about valuation.

To make this clear, consider expected utility theory, which will serve as my initial focus in the presentation that follows. Expected utility theory proposes that choosers should (1) show complete and transitive preferences, and (2) obey a choice separability constraint in a way that seems quite reasonable (amongst other things). Saying that a chooser obeys the axiom of complete and transitive preferences is simply saying formally that she could not be induced to:

1. pay *1 cent and an apple* for an orange, then
2. pay *1 cent and that orange* for a pear, **and**
3. pay *1 cent and that pear* for the original apple.

By the same token, saying that a chooser obeys the separability axiom (a variant of the more widely known independece axiom) is simply the assertion that she cannot:

1. prefer an apple to an orange, **and**
2. prefer 1 cent and an orange to 1 cent and an apple.

This is the reason that these axioms were included in expected utility theory. What is interesting and powerful about the theory, though, is that any chooser who obeys these rules (and the other axioms of the theory) behaves exactly as if she had a stable monotonically increasing utility function and as if her choice behavior was aimed at maximizing her net utility according to that function. Saying someone behaves according to these sensible rules is mathematically equivalent to saying that it looks as if she is trying to maximize some specific utility function. That is an important insight into valuation that neurobiologists cannot afford to ignore[3].

So what are these inferred utility functions like? A subject who behaves according to the axioms behaves "as if" she is maximizing some utility function, but how heavily does even an infinitely large dataset of choices constrain our understanding of this function? To understand the answer to this question, it is

---

[2] And in fairness, I have probably been more guilty of this kind of metaphor than anyone else – a looseness in language that I now regret.

[3] Let me stress here a point that may not be entirely obvious: "utility" really *is* "choice" when these axioms are obeyed. Utility is not "a feeling," or "happiness," or "a hedonic impulse." Utility is a common scale for valuation which gives rise to choice when choice obeys these axioms. If choice obeys these axioms, it is just *as if* a utility function gave rise to these choices. Or, put the other way, if you had a measured function having these properties, and it could perfectly predict choice then it *would be* a utility function and the choices would of necessity obey the axioms. Period.

necessary to understand three important features of utility that have not always been well enough appreciated by neurobiologists.

1. *Utility is ordinal*

Observations of choice constrain the shape, but not the scaling, of the utility function. We might be able to show that a set of observed choices and an assumption that the chooser being studied obey the axioms of the theory are consistent with the idea (to take a simple example) that the utility of money to that chooser could be described as:

$$Utility = Dollars^{0.6}$$

With this equation, we can predict whether that chooser prefers a sure gain of $10 over a 50% chance of $22, and if the assumptions are correct for this chooser then our prediction is also correct. What is interesting to note, though, is that all of our predictions would also be correct if we had written her utility function as:

$$Utility = 50 \times Dollars^{0.6}$$

or

$$Utility = 1000 + Dollars^{0.6}$$

This means that there are multiple equivalent representations of this subject's utility function. We can predict choices by using any of these equivalent utility representations, but the one we employ in a given set of calculations is arbitrary. It is for this reason that economists refer to utility as an *ordinal scale* rather than as a discrete cardinal scale. To make the importance of this insight clear, consider a chooser who prefers apples over oranges and oranges over pears. If we assume the axioms of expected utility theory for this person, we can say that this behavior ranks the utility of the three objects. We can even arbitrarily assign apples a utility of 1 and pears a utility of 0 (for this chooser). Next, we could use lotteries (which of the following do you prefer: *a 50% chance of an apple and a 50% change of a pear* **or** *an orange*?) to place oranges on this same scale, for example at a utility of 0.3. But consider what happens when we suddenly introduce kiwi fruits to the chooser and it turns out that she prefers kiwis to apples. Then the entire scale must be regenerated. This *is* what is meant when an economist writes the word utility and it is no problem mathematically, but it points up an important feature. Utility functions are not cardinal sets of numbers that have definite values that can be added and subtracted. They are ordinally arranged relations between choice objects, and this places important mathematical limits on what you can (and cannot) do with utility functions as objects. The theory of value imposed by expected utility does not include predictions about the cardinal relations of utilities, *by design*. For a neurobiologist, this imposes a particularly burdensome constraint. It implies that is meaningless to say that a neuron's firing rate *is the representation of* utility because such a statement lies outside the domain of expected utility theory. Neurons yield to us a fully cardinal measurement when we observe their firing rates. Firing rates are numbers that can be added and subtracted in a way that utilities cannot. If a neuron had a firing rate that revealed the desirability of an apple irrespective of the other objects placed before that chooser (as in the study of Padoa-Schioppa and Assad, 2008), that neuron could *not* be said to encode the utility of apples, because one feature of this powerful theory is that the utility of apples is not a unique number[4]. Of course that firing rate could be linearly proportional to utility. If we increase the number of apples presented to the subject until the firing rate doubles we might be able to conclude that utility has doubled, but utility and firing rate would remain distinct in this important way.

2. *The Axioms of Utility Theory are not Always Consistent with Choice*

Humans do, on occasion, both prefer an apple to an orange and also prefer 1 cent and an orange to 1 cent and an apple. This is (speaking a bit imprecisely) what the Allais (1954) paradox shows (see Chapters 1 and 11 for more on this paradox). Of course, this has implications for economics, but it also has a huge implication for neuroscientists. If a neuron had a firing rate that was always linearly proportional to utility, then the firing rate of that neuron *could not be used to always predict real human choice*. A neuron with a firing rate proportional to utility would – by definition – obey all of the axioms of expected utility. It could not generate the Allais (1953) paradox, because that is the nature of what is meant by *utility*.

3. *Utility Implies Agent Welfare/Agent Well-being*

One of the most important functions of economics is to tell us whether a change in policy or government will make people better off. Economists have often argued that expected utility theory helps them make

---

[4] It is important to point out that this is not a limitation only of expected utility theory; it is a feature of almost all economic theories of value.

this determination. If, for example, all citizens were to obey the axioms of expected utility theory, then governments would have an easy time keeping them happy or, more formally, maximizing their welfare. Since a chooser who obeys expected utility always acts to maximize her own utility, then we can maximize her welfare by allowing her the freedom to choose whatever she desires. This is a common (though not necessarily ubiquitous) approach to figuring out how to design policies that maximize the well-being (or technically the *welfare*) of individuals.

It is important that neuroeconomists be aware of this issue, because if a neurobiologist was to argue that the firing rate of a neuron was linearly proportional to utility he might be heard as saying that maximizing the firing rates of those neurons in citizens would be maximizing their welfare, even if the firing rate of that neuron *could not be used to predict real human choice*. I think that neuroscientists need to very carefully avoid making such a claim for the foreseeable future. In almost all of the neuroeconomics studied to date, we have tried to link activity in the nervous system to choice behavior. This volume shows how much we know about the neural circuits that give rise to choice. By contrast, we know very little today about the neural circuits that give rise to an individual's sense of well-being. If we did understand those neural circuits, then we might be able to make some claims related to welfare issues in economics. In any case, we do not have such expertise at this time, and I want to take care to emphasize that it is the neural mechanism of choice and not the neural mechanism for experiencing well-being that this standard back-pocket model attempts to describe.

Defining Subjective Value (SV)

The task before us is to ask, how can we relate neuronal firing rates, or measurements of the BOLD signal, to the valuations of actions and objects that we believe guide behavior? One way to proceed is to try to relate these activation patterns to expected utility. Under many conditions expected utility theory does predict choice, and that seems to be an observation that we do not want to overlook. On the other hand, one of the reasons that we want to develop an algorithmic model of decision making is that we have every reason to believe that such an algorithmic model would predict choice behavior even when expected utility theory cannot. So how do we gain access to the theoretical power of expected utility theory without becoming burdened with its failures, and in a way that respects the two-stage model for valuation and choice that is developing today? One has to note here that for many economists this is a critical point – and one about which there has been much confusion. To resolve this confusion, I suggest the following definition:

> **Subjective value**: Subjective values, at least for the purposes of this initial argument, are real numbers ranging from 0 to 1000. They take as their natural units action potentials per second. Subjective values have the following properties:
>
> 1. Subjective values are equal to (or better yet defined as) the mean firing rates of specific populations of neurons, the identification of which follows. For this reason, subjective values are linearly proportional to the BOLD[5] signal as measured in these same populations.
> 2. Subjective values predict choice stochastically. More formally, I define them as the sum of true subjective value and a noise term (see below). This means that subjective value theory will be most closely allied with random utility-type models from economics.
> 3. When expected utilities predict choice behavior, subjective values are linearly proportional to those expected utilities.
> 4. Subjective values are *always* consistent with choice, though stochastically, even when choice is not consistent with expected utility theory.
> 5. Subjective values have a unique *reference-dependent* anchoring point called the baseline firing rate. All subjective values are encoded cardinally in firing rates relative to this baseline. This means that subjective value theory will be most closely allied with reference dependent forms of utility-type models from economics.

Of course I recognize that some of these properties will have to be relaxed, but probably not in important ways. The BOLD signal and mean firing rates, for example, are not exactly linear in their relation, but these five statements capture the central features of subjective value around which our definition of the choice architecture will be organized.

First and foremost, the definition I suggest here allows us to be clear about why expected utility theory will be enormously valuable to the neuroeconomic enterprise. Expected utility theory provides a compact definition that, under at least some circumstances, describes patterns of choices. *Where that is true*, measurements of utilities tell us unambiguously what SV *must* look like to within a linear transformation. (This, of course, assumes that we can find a mean neuronal firing rate that looks like a candidate for encoding SV, but we turn to that in the next section. For now, we simply seek clear definitions of our goals.) Second, this definition says that if we could actually measure SV, we would be able to use those measurements to predict behavioral violations of expected utility theory, like the Allais paradox, as well as human

---

[5] The Blood Oxygen Level Dependent Signal of functional Magnetic Resonance Imaging (fMRI).

choice behavior well captured by other approaches like prospect theory. Third, SV must be subject-specific. This follows, of course, from its relation to the utilities of expected utility theory.

SV is defined in units of average spikes per second as an object that predicts the choices of individuals. Finally, I want to be clear that measurements of SV *do not* have clear welfare implications. Because SV does not (at the very least) obey the independence axiom globally (since human choice does not obey this axiom), maximizing SV will not yield a maximization of something like a complete and transitive preference function. Further, and probably more importantly, SV predicts choice. Because we are modeling at an algorithmic level, this does not mean, *ex ante*, that SV is related to a chooser's sense of well-being. That may be mediated by other neural systems. Welfare maximization and SV maximization should not be equated.

To summarize, I define here the concept of subjective value which is meant to be a fully cardinal object with several important restrictions. At least initially, it cannot take negative values (an important point to which we will return). It has both a finite range and finite (and large) variance. The importance of this point will be immediately clear to neurobiologists. For economists, it means that errors and stochasticity in choice are unavoidable features of the architecture. This suggests properties related both to random utility models and to stochastic errors in choice. The importance of this point will be immediately clear to economists. For neurobiologists, it means that whenever choice behavior obeys the axioms of random utility models we know a tremendous amount about how a final common valuation system ought to be behaving. Following this line of reasoning, then, my hypothesis is that SV is encoded directly in the valuation mechanisms of the human brain and that existing economic theory tells us much about how this representation must behave. As we make measurements to prove this, we will be able to place additional important restrictions on SV.

Finally, I need to make it clear that what I am suggesting is that one central goal of neuroeconomics should be to develop a complete theory of SV. As that theory is enriched, it will continue to refine our understanding of which economic theories are better than others at predicting SV. Random utility models, for example, will be shown below to be better predictors of SV than traditional utility models. Reference-dependent utility models will also be shown to be better predictors of SV than traditional consumption utility models. Whether traditional economists will care that empirical constraints on SV can be used to identify some utility-based models as closer fits to the human choice architecture will, of course, be a matter of taste.

Relative Subjective Value (RSV)

We define the relative subjective value of a single option *j* as:

$$RSV_j = \frac{SV_j}{\sum SV_i + c}$$

where *i* is the set of all options in a choice set (including *j*) and c is an empirically measurable normalization constant of the type first described in the cerebral cortex by Heeger (1992). Our current evidence suggests that choices are actually made between options by comparing RSVs after corruption by noise. The evidence for this arises from work in parietal cortex, which is summarized in Chapters 4, 28, 29, and 31.

Obtained Subjective Value (ExperSV)

Obtained Subjective Value is a pattern of neuronal firing, much like SV, which encodes the subjective value of current states of the world[6]. The neural location of ExperSV is not known, though the activity of dopamine neurons provides overwhelming evidence that it is present as one of the midbrain inputs to those neurons. For reasons that will be described below, Obtained Subjective Value actually serves as one source of the utility-like properties of SV.

Reward Prediction Error (RPE)

RPE is defined here as in learning studies and as summarized in Chapter 22. It is:

$$RPE = \alpha(SV_{forecast} - ExperSV)$$

Stochastic Terms

The existing neural data suggest two sources of stochasticity that influence choice: one at the level of the valuation system and one at the level of the choice system. In economic terms, the first can be viewed

---

[6] Let me stress again that ObtainedSV is not necessarily a welfare measurement.

as roughly corresponding to random utility distributions[7] (McFadden, 1974; Gul and Pesendorfer, 2008) and the second as corresponding to the Trembling Hand[8] (Selten, 1975) notion of stochastic behavior. There is compelling evidence for both such sources.

Subjective value variability is a random term drawn from distribution assumed to be Gaussian and added to mean *subjective value* to yield SV. It is always present. The variance of this term may or may not be adjustable.

Cortical noise before choice is a final noise source added (as a stochastic time series) to RSV before choice occurs. The source of this term is noise intrinsic to cortical neurons, which requires that it be Poisson in distribution at the mechanistic point of addition (see, for example, Tolhurst *et al*., 1983). Neuronal pooling that occurs during the choice process, and adjustability of the inter-neuronal correlation term, may be used to reduce this variance (Krug *et al*., 2004). For more on the theoretical implications of this, see Glimcher (2005).

Valuation Mechanisms and Subjective Value

Formally (and of course too simplistically), subjective value (and terms that inherit properties from subjective value) can be seen as a neuronal sum of the form:

$$SV_j = \frac{\sum_i \omega_i x_{ij}}{\sum_i \omega_i}$$

where the term i indexes each of the neurons in the brain, x_i is the firing rate of the ith neuron, and ω_i is a weight ranging from 0 to 1 describing the additive contribution of that neuron to the SV of object or action j. This object places into the language of economics the standard neurobiological insight that a weighted sum of neurons in topographic maps encodes behaviorally relevant variables (for more details on this neurobiological issue, see Lee *et al*., 1988.) The subjective value of a particular object in the external world, j, is thus simply represented as the average weighted firing rate of a subpopulation of neurons that encodes the subjective value of that object[9]. In a topographically mapped action-encoding region like the superior colliculus, this is equivalent to saying that activity in a restricted region of the map encodes value for a particular action. I should note, however, that this definition specifically excludes distributed non-linear encoding schemes[10]

For an empirical neurophysiologist or functional magnetic resonance imager looking for SV_j in the brain, two questions then become paramount:

1. Is there a firing rate pattern (or a BOLD activation in the case of fMRI) we can identify in the brain that is linearly correlated with the utility of actions or objects (when utility predicts choice)?
2. What is the most compact population of neurons (both in number of neurons and in anatomical extent of the population) that can maintain this linear correlation with SV_j (i.e. the smallest population of neurons for which ω_i is not equal to zero).

The data we have available today suggest that two brain areas seem likely to contain all the neurons we require to extract SV for any object: the ventral striatum and the medial prefrontal cortex.

THE BASIC STRUCTURE OF THE VALUATION SYSTEM

If one accepts that (1) mammals evolved to learn the values of different states of the world both when their actions influenced those states and when they did not, and (2) that choices (the selection amongst available options) must be complete in the nervous system before actions can be planned and executed, then one can hypothesize that valuation must be at least partially autonomous of choice and that the process of choice must be complete before action is produced. These hypotheses seem to have been validated by a wealth of empirical research in neuroscience conducted over the past couple of decades. Indeed, amongst neurobiologists there is essentially universal agreement that a group of neural systems for valuation has been identified.

---

[7] Note that random utility theories form a class of economic models in which it is assumed that the subjective valuations of options vary stochastically, but that choosers always select the current best option from that stochastically varying set. In these models, it is the perception of value itself that is hypothesized to vary. Perhaps surprisingly, these models place some very interesting constraints on the relationship between value representations and choice.

[8] In contrast, trembling hand models propose that stochasticity in choice arises from errors during the choice process which lead to the selection of suboptimal elements from the choice set. These models place other interesting constraints on choice. One interesting signature of models of this type is a dependency of errors on choice set construction.

[9] This of necessity excludes non-linear interactions like those encountered in a game theoretic specification of SV. While it is not necessarily my intent to exclude these other kinds of interactions, the available data suggest that SV actually is linear with these firing rates. In any case, this definition could be relaxed.

[10] A constraint that could, at a later date, also be relaxed.

Almost certainly, the critical historical event that pointed towards a common neural system for valuation was the study of learning and dopamine, a topic reviewed in detail in Part 3 of this volume. Understanding why the study of dopamine led to postulates about valuation, however, requires an examination of the history of dopamine. In the 1920s, the German physiologist Otto Loewi established that neurons communicated with each other through a chemical mechanism we now call neurotransmission. His groundbreaking work established the existence, however, of only one neurotransmitter through which it was believed all neurons communicated. The existence of multiple neurotransmissive systems was revealed in 1964, when Dahlström and Fuxe (1964) visualized adrenaline-, noradrenaline-, and dopamine-containing neurons. These measurements revealed a set of anatomically and neurochemically discrete brain systems, a structural feature of the nervous system that had not been previously identified. Of particular interest, for our purposes, was the discovery of two to three groups of dopamine containing cell bodies that projected from the midbrain (from two areas called the *substantia nigra pars compacta*, *SNc*, and the *ventral tegmental area*, *VTA*) to the basal ganglia and the frontal cortex (Figure 32.2). By the 1970s and 1980s it had become clear that many drugs of abuse acted through this system, suggesting a role for dopamine in hedonic experience.

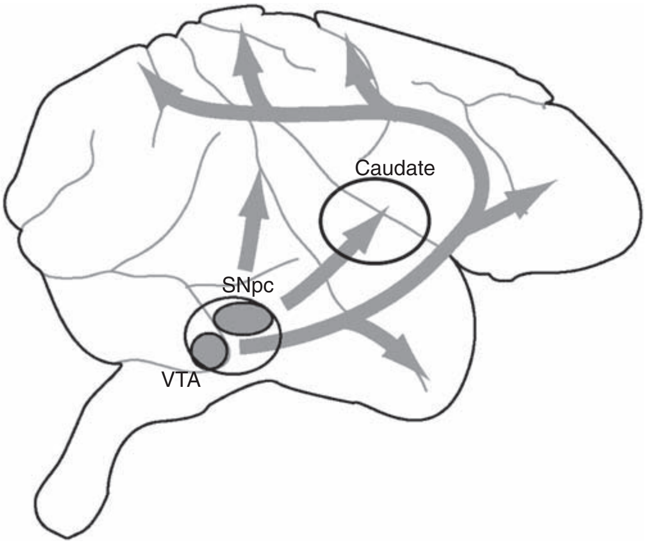


FIGURE 32.2 The principal dopaminergic pathways of the midbrain. SNpc, substantia nigra pars compacta; VTA, ventral tegmental area.

Learning Subjective Values

The critical breakthrough that allowed modern studies of valuation to crystallize around the midbrain dopaminergic pathways, however, was the work of Schultz and colleagues (1993). These authors measured the spiking activity of single dopamine neurons while monkeys passively received rewards during a classical conditioning task (see Chapter 22 for more details). They found that *unconditioned* rewards produced a strong response in these neurons, while conditioned rewards did not. This was an important finding, because it revealed that the activity of dopamine neurons *could not* simply code hedonic experience. This led Montague and colleagues (1997) to propose that dopamine neurons encoded the difference between expected and obtained rewards; the *reward-prediction error* of learning theory. What followed were a host of papers that established that dopamine firing rates could be described as:

$$DA(spks/s) = \alpha(ExperiencedR - ExpectedR)$$

where DA is the instantaneous firing rate of VTA and SNc dopamine neurons, *ExpectedR* is the magnitude of the reward expected by the subject at this time, *ExperiencedR* is the magnitude of the reward being experienced by the subject at this time, and α is a scaling parameter that controls the subject's learning rate as described below and in Chapter 22. If ExperiencedR is reward value in an arbitrary currency and some other computational element simply recomputes after every expected or unexpected reward

$$Q_{j,t} = Q_{j,t-1} + DA$$

where j indexes actions or goods and t indexes time, then Q represents a current estimate of the expected value of action or good j. For economists, I note this is a recursive form of the Bellman equation computing a reversed discount function in which α describes a *forgetting rate*.

A very interesting feature of this system, which is, however, often overlooked, is the units in which *ExperiencedR* encodes the magnitudes of rewards being received. Consider a primary reward like water. If *ExperiencedR* encodes water in milliliters, then Q_j converges towards a representation of expected value in milliliters. If, however, *ExperiencedR* encodes something like the utility of water, or more formally it encodes *ObtainedSV*, then the system converges not towards a representation of expected value but towards a representation of SV. This point is critical because neural systems of valuation must estimate real-world values through the sensory responses of neural transducers, and any upstream transformation of information by these transducers is propagated through the system – thus, if volume encoding afferents from the tongue, like all sensory afferents that have ever been studied, encode a Stevens Power Law (Stevens, 1970) compressed representation of volume, then the set of Q_js computed from dopaminergic activity would encode (or inherit) a power function representation of magnitude. It seems almost certain that this has to be the case, given that every sensory system ever studied employs a power law for compression (see, for example, Stevens, 1970). This leads to speculation that the encoding mechanism for primary rewards serves as at least one (if not the) source for the curvature of the utility function amongst primary rewards. It is also a possibility that this may offer some insight into the sources of reference dependence in human choice behavior.

For these reasons, it is tempting to speculate that dopamine neurons should receive as an input *ObtainedSV* from which their target neurons learn SV directly, thus accounting for the source of curvature in the utility functions for primary rewards with the neural hardware for sensory encoding. Of course this places some interesting constraints on how the DA neurons should work. They cannot, for example, code RSV (relative subjective value) with regard to a finite choice set, because if they did then SV could not be computed from them. The pre-existing body of economic theory makes this clear. If they did code RSV, the stored SV of a good or action would be dependent on the choice set within which it was learned. As a result, the choice mechanism would be unable to obey the axioms of complete and transitive preference which both humans and animals often obey.

This is an observation, however, that may seem at first blush to contradict data in the literature (Tobler *et al.*, 2005), and this apparent contradiction is important because it highlights the power of economic theory in neuroscience. These data suggest that the RPE signal measured in the dopamine neurons is variance dependent. As the variance of the reward stream increases the magnitude of the dopamine firing rate for a given $ExpectedR - ExperiencedR$ goes down. Theory tells us, however, of the importance of complete and transitive preferences and what they imply for valuation mechanisms, and seems to suggest that $ExpectedR - ExperiencedR$ cannot scale with the variance of the choice set and still preserve transitivity in the stored representation of value. The resolution of this apparent paradox is that it must be the learning rate, and not $ExpectedR - ExperiencedR$ itself, which scales with variance. In other words, we can say that if subjects are transitive behaviorally, we can reject the hypothesis that $ExpectedR - ExperiencedR$ scales with variance[11]. Our ability to make this statement comes from the explicit linkage of theory and measurement. Of course, this also raises the possibility that violations of complete and transitive preference – when these violations do occur – may reflect features of this measurable set of computations.

In any case, the dopamine neurons broadcast this signal throughout the frontal cortex and basal ganglia – also suggesting a role for all of these areas in learning. The observation that the firing rates of dopamine neurons encode the difference between expected and obtained reward is critical, because it reveals that the inputs to the dopamine neurons include a signal both of both the value of the reward that was received and the value of the reward that was expected – clear evidence that a valuation signal of some kind, an object we have defined as *ObtainedSV*, must exist[12].

Though most of the work described above was conducted in animals, there is clear evidence that these dopaminergic neurons behave in the same manner in humans as they do in all other living mammals. Like other mammals, humans find dopaminergic drugs reinforcing. Like other mammals, humans have these same dopaminergic pathways. Like other mammals, dopaminergic drugs can be shown to bind to receptors in the terminal fields of these neurons. But the best evidence for the notion that a circumscribed learning-based valuation system associated with dopamine occurs in humans comes from fMRI studies of humans engaged in learning about rewards. In 2002, two groups (O'Doherty *et al.*, 2002; Pagnoni *et al.*, 2002) demonstrated simultaneously that activity in the dopaminergic terminal fields of the striatum and the frontal cortex during both gustatory and

---

[11] In fact, this scaling of the learning rate with variance is a feature of efficient learning systems – a fact well described in any mathematical treatment of the Kalman filter.

[12] To be more precise, there is now some compelling evidence that dopamine firing rates encode only positively valued, or near positively valued, reward-prediction errors (Bayer and Glimcher, 2005). There have been hints of this in the literature for some time (Hollerman and Schultz, 1998). The suggestion here is that positive and negative RPEs may be encoded separately in the nervous system (Daw *et al.*, 2002). The idea of splitting the RPE term into negative and positive elements should be naturally attractive to behavioral economists. We have known since the work of Kahneman and Tversky (see, for example, Kahneman and Tversky, 1979; also Chapter 11 of this volume) that human choosers are more sensitive to losses than to gains – a feature known as loss aversion. If positive and negative RPEs are coded by different systems and those systems map positive and negative values of $ExpectedR - ExperiencedR$ to firing rate with different gain terms, then the ratio of these two independent gain terms could well account for some features of loss aversion.

monetary reward tasks behaved exactly as predicted. This indicated, basically beyond a reasonable doubt, that there existed a valuation-learning system in the striatum and frontal cortex of humans. So to summarize, this leads me to suggest that dopamine neurons lead to the direct computation of SV under some conditions:

$$SV_{jt} = SV_{j(t-1)} + \alpha(ExperSV - SV_{j(t-1)})$$

where SV_{jt} is the subjective value of object or good j, which is learned from repeated experience, as estimated at time t. Note that, as mentioned above, *ExperSV* for primary rewards is a compressive function of *ExperiencedR*, as is really required by what we know of sensory encoding systems. This means that risk aversion, at least over primary rewards, is the product of Weber-type encoding mechanisms in our sensory apparatus.

What remains, then, is to understand where and how SV is mechanistically computed and stored. Two lines of evidence contribute to our understanding of these issues: neuronal recording studies in animals and fMRI studies in humans. The recording studies in animals have now established that the basal ganglia contain essentially all of the computational elements required for the execution of reinforcement learning (or, more precisely, *temporal difference* learning) algorithms. There are, for example, neurons within the basal ganglia that encode the magnitude of reward that an animal expects to receive for producing a particular behavioral action, neurons that encode the actions that have just been executed, and neurons with firing rates dependent on the current state of the environment, amongst other things. These neurons are located in the striatum and project out of the basal ganglia largely through the ventrolateral nucleus of the thalamus, which in turn projects back to the frontal cortex. Single unit recording studies in the frontal cortex have also demonstrated the existence of neurons that encode values, but this time the values of goods, not of actions (see Chapter 29). fMRI studies in humans tell a similar story (see Chapters 23 and 24 for more details), suggesting that frontal and basal ganglia circuits form the core of the human mechanism for RPE-based value learning.

There is, however, evidence for other learning mechanisms in these same structures which interact with this well studied RPE-style learning mechanism. The details of these other learning systems are still being worked out, but what is known to date is described in Chapters 12 and 24 of this volume. In essence, these studies suggest that a set of mechanisms, most if not all interacting with dopamine, provide tools for learning and representing value in the frontal cortex and the basal ganglia.

For a neuroeconomist, then, these studies constitute overwhelming evidence that a value system exists and can be functionally localized. Where, then, is the final point of convergence at which SVs are passed to the choice system? Put more formally, in a preceding section I argued that subjective value can be seen as a neuronal sum of the form

$$SV_j = \frac{\sum_i \omega_i x_{ij}}{\sum_i \omega_i}$$

where the term i indexes each of the neurons in the brain, x_i is the firing rate of the ith neuron, and ω_i is a weight ranging from 0 to 1 describing the additive contribution of that neuron to the SV of object or action j. The question we need to answer is whether there is an anatomically discrete neuronal population that can supply all the non-zero values for ω required by the choice system.

One way to begin to answer this question is to look at the existing fMRI data and ask: is there a small number of areas that are actively correlated with SV under essentially all reward and choice conditions that have ever been studied? Perhaps surprisingly, the answer to this question seems to be yes: the ventral striatum and the medial prefrontal cortex show up in dozens of studies under essentially all choice conditions as coding something like SV.

Activity in the ventral striatum has been shown to be correlated with both rewards and punishments (Delgado *et al.*, 2000), the magnitude of cumulative rewards (Elliot *et al.*, 2000), the anticipation of reward (Knutson, 2000; Knutson *et al.*, 2003), the expectation of monetary reward (Breiter *et al.*, 2001), the expectation of primary rewards (O'Doherty *et al.*, 2002), the receipt of monetary rewards (Elliott *et al.*, 2003), monetary expected values (Knutson *et al.*, 2005), behavioral preference rankings amongst rewards (O'Doherty *et al.*, 2006), potential gain magnitude and loss magnitude as scaled by subject-specific levels of loss aversion (Tom *et al.*, 2007), and discounted reward value at delays ranging from minutes to 6 months (Kable and Glimcher, 2007). Single unit recording studies of the dorsal striata of monkeys, both in the caudate (Lau and Glimcher, 2006) and in the putamen (Samejima, *et al.*, 2005), tell a similar story. Neurons in these areas have been identified which code action values. All of these data suggest that whenever rewards are received or preferences are expressed, activity in the ventral striatum encodes the magnitudes of those rewards or preferences[13].

A similar story seems to hold in the medial prefrontal cortex. Activity in this area has been shown to be correlated with monetary reward magnitude (Knutson *et al.*, 2000, 2003), preference ordering amongst primary rewards (McClure *et al.*, 2004a), the expected value of a lottery (Knutson, *et al.*, 2005), the subject-specific valuation of gains and losses (Tom *et al.*, 2007), the subject-specific discounted reward value (Kable and Glimcher, 2007), and willingness to pay (Plassman *et al.*, 2007). Activity in this area appears to be correlated with valuation under all of these conditions.

This leads me to propose that mean activity in the medial prefrontal cortex and the ventral striatum encodes SV. Different neuronal subpopulations in these areas must encode different options, and so these areas must employ a complex topographic encoding scheme which segregates the populations that encode the SVs of different actions or goods. The details of this encoding scheme, which probably lies beneath the resolution of fMRI, are only just beginning to be understood, and the encoding schemes employed by the two areas are almost certainly different. The medial prefrontal cortex, because of its closer relationship to goods-related encoding areas like the orbitofrontal cortex, may well encode SV in terms of goods, while the VS may employ an action-based encoding scheme. But in any case, I propose that these two areas serve as the final common representation of SV for use by the choice mechanism.

To be quite precise, I propose that the mean activity in subpopulations of the medial prefrontal cortex and the ventral striatum encodes SV when options are under consideration for choice or the objects of current learning. It is this activity which, I argue, both guides choice and encodes the reward prediction (SV) that is used in learning (probably as resident in the ventral striatum). This information, I suggest, is stored throughout a much larger network of areas spanning the frontal cortex and the basal ganglia in the synaptic strengths connecting neurons in these areas, the strengths of these synapses being set by the well understood biophysical mechanisms of dopamine-dependent long-term potentiation and long-term depression (and perhaps based on the actions of other plasticity generating neurotransmitters like serotonin). When instantaneous subjective value is represented, it reflects, I propose, the sum of activity passing through these synapses located in areas including the inferior frontal sulcus, the insula, the amygdala, the posterior cingulate, the superior temporal sulcus, the caudate, the putamen, and the dorsolateral prefrontal cortex, and impinging on the ventral striatum and the medial prefrontal cortex.

What we know about the biophysics that would be required to instantiate this process have two important implications that need to be mentioned. First, recall that all neurons have a limited dynamic range and a significant finite level of stochasticity in firing rate. This means that instantaneous SV is necessarily drawn at each instant from an underlying distribution. This therefore requires that the notion of SV be closely related to random utility models (and not traditional von Neumann-Morgenstern utility) from economics. Second, it needs to be noted that all neurons have a "baseline" firing rate, and neurons in these areas are no exception. Recently, Tom and colleagues (2007) have shown that activation in these areas continuously represents gains and losses on a common scale with an inflection point at a zero-gain point in those experiments. This suggests that baseline spike rate in these populations is the unique representation of the reference point for SV's reference-dependent encoding of value. Of course this conclusion, even if correct, does not constitute a theory of the reference point; it simply identifies an empirical technique for direct measurement of the reference point.

CHOICE

Unlike valuation, which has been extensively studied in both humans and other animals, choice has been the subject of study principally in awake-behaving monkeys in neuroscience. This may reflect the fact that the temporal dynamics of choice make it difficult to study with fMRI. In any case, an understanding of choice requires an understanding of existing work in non-human primates.

Initial studies of choice in monkeys evolved almost simultaneously from studies of sensory-perceptual systems (see, for example, Newsome *et al.*, 1989) and movement control studies (e.g. Glimcher and Sparks, 1992). The most important of these studies examined how monkeys used noisy visual-sensory signals to identify one of two orienting eye movements, or saccades, as reinforced. They did this by leveraging an extensive pre-existing literature on the structure of the visual and eye-movement systems to search for the decision-making circuits which connected them in

---

[13] However, one possibility that needs to be ruled out is that activity in the ventral striatum encodes only reward-prediction errors and not SV *per se*. Available single unit data rule this out in all other areas of the striatum, but the definitive study has not yet been conducted in the ventral striatum.

these tasks (details of this line of study can be found in Chapters 4, 28, 29, and 31 of this volume). Subsequent work has generalized many, but not all, of these findings to arm-movement control systems and to studies of humans.

We have to begin, therefore, with a review of the basic structure of the saccadic control system (Figure 32.3). The lateral intraparietal area (LIP) in the posterior parietal cortex is one of the critical elements in this system, and it consists of a roughly topographic map both of objects in the visual world and the eye movements that would required to align gaze with those objects (for a review, see Glimcher, 2003). Thus, a particular location on the map (or more precisely the neurons on the map at that location) activates when a visual stimulus appears 10° to the right of fixation, *and* that region might become particularly active milliseconds before an eye movement which shifts gaze 10° to the right. This area, in turn, projects both to the frontal eye-fields and the midbrain superior colliculus, two additional topographic maps that are broadly similar in function. The frontal eye-fields project, as well, to the superior colliculus directly. A final note is that many of these areas are reciprocally connected (for a review of this anatomy see Platt *et al.*, 2003), a fact which is probably important for understanding choice. Finally, the colliculus is connected to brainstem circuits that actually govern eye movements in real time. The connection between these brainstem systems and the colliculus are mediated by a class of collicular neurons called "burst" neurons. Burst neurons have the interesting biophysical property that they can fire action potentials in either of two states: a continuous low-frequency state in which many different firing rates are observed, and a burst state characterized by a fixed and extremely high firing rate.

It is widely assumed that actual generation of a movement involves driving the collicular burst neurons above a specific firing-rate threshold, after which a burst occurs that is self-perpetuating and persists until the movement is complete. Inhibitory interconnections in the collicular map seem to preclude burst-like activity occurring at more than one location at a time, suggesting that the collicular architecture allows only a single movement to be executed at a time. Studies in area LIP, the frontal eye-fields, and the superior colliculus all indicate that low-frequency firing in all three is related to the probability that a movement will be executed by the animal. To be more specific, if a particular movement is likely to yield a reward, then activity in all three maps at the locations associated with that movement is elevated. Of these three maps, the one that has been most studied with regard to decision is LIP. In LIP, it has been shown that if the magnitude of a reward or the likelihood of a reward is systematically manipulated, then firing rates in these areas are a roughly linear function of those variables under many conditions. To be yet more precise, current data suggest that activity in this map encodes relative subjective value (RSV) of the type defined above.

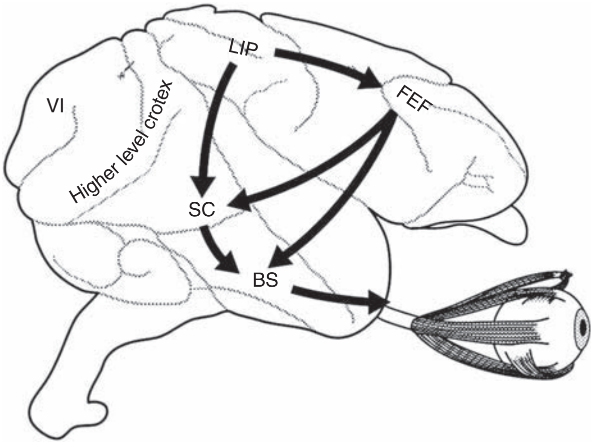


FIGURE 32.3 The saccadic control system of the rhesus monkey in which most studies of the choice mechanism have been studied. V1, primary visual cortex; LIP, lateral intraparietal area; FEF, frontal eye fields; SC, superior colliculus; BS, brainstem eye movement control circuits.

Together, these data suggest the following model for eye-movement generation. At any moment in time, neurons in LIP represent the instantaneous RSV of each movement in the saccadic repertoire. Movements that have non-zero values are thus each represented by local activity on the map that is linearly proportion to RSV (see, for example, Dorris and Glimcher, 2004). I hypothesize that the representation of SV localized in the medial prefrontal cortex and the ventral striatum serve as the initial source of this signal. Previous studies have noted that other (and perhaps all) cortical areas perform a divisive normalization on their input data (Heeger, 1992; Schwartz and Simoncelli, 2001). It has now been observed that (at least to a first approximation) this also occurs in area LIP (Dorris and Glimcher, 2004; Sugrue *et al.*, 2004; Louie *et al.*, 2007), and the result is likely a shift from SV to RSV in the posterior parietal cortex.

RSV, it should be noted, would serve to map SV into the limited dynamic range of the LIP neurons. LIP neurons are limited in number, fire over a roughly 100-Hz dynamic range, and have (errors that are drawn from a) Poisson-like distribution. This means that the representation of RSV, rather than SV, in this structure may solve an important problem. The shift to RSV guarantees a distribution of the SVs of the current choice-set over the limited dynamic range of these neurons. Unfortunately, the finite dynamic range and noise associated with these neurons may also impose a constraint. As the choice set becomes larger, noise may swamp the signal, leading to profound inefficiencies when selecting amongst large numbers of possible movements. One is tempted to speculate that this may, in fact, be a neural account for the choice-set size effects which have recently been examined in human choosers (see, for example, Iyengar and Lepper, 2000). It may also be that processes like choice set editing are tools used by frontal areas to winnow the size of the choice set operated on in parietal areas.

It should also be noted that the Poisson variance of these neurons may serve a useful function by allowing for stochasticity in behavior under conditions in which behaviors like mixed strategy equilibria arise. It was Nash who noted that mixed strategy equilibria arise when the expected utilities of the strategies being mixed are equivalent. In a similar way, when the RSVs of two options are equivalent it might be expected that the stochastic nature of these neurons yields mixed-strategy behavior. If these neurons are always stochastic in their behavior – a hypothesis that has been largely documented in monkeys – patterns of activity in LIP may be related to economic notions of the trembling hand (Selten, 1975).

In summary, then, the available data suggest that at all three of these areas – LIP, FEF, and SC – carry signals encoding RSV, and that movements occur when activity associated with one of the positively-valued options drives its associated collicular neurons into their burst mode. A tremendous amount of work (again summarized in Chapters 4, 28, 29, and 31) has examined this process of movement-triggering under conditions in which animals are instructed to make movements as quickly as possible. Less is known about how movement selection is triggered in non-reaction time settings. One important possibility is that an input to one or more of these areas alters the inhibitory interactions within the map, forcing convergence to a single action.

The basic model proposed for selecting eye movements is thus that signals encoding SV project to these areas, probably through LIP, which normalizes those signals to represent RSV which is further contaminated by local noise, the degree of which across the entire population may be regulated by adjustable inter-neuronal correlations (Glimcher, 2005). These signals propagate recursively through these networks while reflecting SV inputs that may be entering the maps at many locations. An external signal then permits, or forces, convergence of the network to a single choice which occurs when the collicular neurons are driven above their burst threshold.

Preliminary evidence for this hypothesis has been gathered by Louie and Glimcher (2006), who have shown that early in a trial the neurons of LIP represent the RSV of discounted gains associated with specific saccades, and that it is only later in the trial that these same neurons come to encode information about the actual choice made by the animal. This seems to suggest that the basic model is sound, at least for tasks of this type.

Two questions, however, immediately arise: how does this system achieve choice amongst more abstract objects that do not have specific movements associated with them, and does this model generalize to humans and non-eye movement conditions? A limited amount of data exists which suggests that this general class of system does operate under conditions in which choices are made between more abstract objects. Gold and Shadlen, for example, demonstrated that when animals must choose between red and green targets that constantly interchange locations, activity in the superior colliculus reflects the instantaneous mapping between color and value even if this changes from trial to trial (Gold and Shadlen, 2000; see also Horwitz and Newsome, 2001; Sugrue *et al.*, 2004). This clearly indicates that the saccadic choice circuit has access to instantaneous mapping information relating abstract properties to actions. It cannot tell us however, how choice is accomplished (or if it can be accomplished) in the absence of any mapping to motor circuitry of any kind.

We do, however, have some interesting hints that these choice circuits are interconnected with important valuation areas in the frontal cortex and basal ganglia. Padoa-Schioppa and Assad (2006), for example, have demonstrated the existence of neurons in the orbitofrontal cortex that encode an animal's choice before the movement expressing that choice is executed. In a similar way, Lau and Glimcher (2006) have observed choice neurons in the dorsal striatum. At the very least, this suggests that the choice circuit can send information about decisions frontally, but it may also indicate that these areas participate directly in the convergence process by which choice is accomplished.

The question of whether these circuits that have been so well studied in monkeys can be generalized to other classes of movements and other species is one about which we have much less information. We do know that adjacent to area LIP are areas specialized for arm, hand, and face movements. Standard theories suggest that a group of areas lining the intraparietal sulcus serve as movement-control interfaces for all of the body, although there are problems still being resolved with those hypotheses (*cf.* Levy *et al.*, 2007).

But it does seem clear that the general theories of movement control advanced for the oculomotor system do have analogues in the skeletomuscular system. Further, injuries to any of these systems, in either humans or monkeys, leads to permanent deficits not in the musculature but in the ability to produce movements. Finally, a small number of fMRI studies have shown value-related signals in the posterior parietal cortex, although these signals are almost always of weaker magnitude than in more frontal areas. This, of course, raises the possibility that the weaker fMRI signal reflects the temporal dynamics of choice observed in the Louie and Glimcher (2006) study. Because subjective value is only represented until a decision is made, in these areas the magnitude of the SV signal, integrated over an entire trial, may be much less than in areas located more frontally where SV is represented throughout a trial.

ALTERNATIVES TO THE TWO-STAGE MODEL

Choice Probabilities

Some early models of the primate choice system proposed that when choosing between two actions a choice probability was computed directly from the identity of the option pair, rather than by comparing something like the utilities of the two options under consideration. This choice probability was then proposed to stochastically direct action. When these models were introduced, some argued that they could serve as an alternative to preference-based models.

Two factors argue against models of this type. The first is axiomatic. Consider an agent who has been asked repeatedly to choose between chocolate and apples. Then she is asked to choose repeatedly between apples and crackers. We can, of course, represent the behavior of the agent with two choice probabilities. If we begin, however, with the assumption that the chooser represents *only* choice probabilities, then we must necessarily remain agnostic about what the agent will select if we offer her a choice between chocolate and crackers. If, on the other hand, we hold a belief that knowing her choices under these first two conditions reveals her likely choice under the third condition, then we are basically assuming complete and transitive preferences that invoke a utility-like representation. In other words, we invoke a system which behaves "as if" abstract valuations, subjective values, are represented.

The second factor arguing against this possibility is empirical. We now have compelling neurobiological evidence that subjective values of some kind are represented in the brains of monkeys (Dorris and Glimcher, 2004; Sugrue *et al*., 2004, Louie and Glimcher, 2006; Padoa-Schioppa and Assad, 2008). In those experiments and others like them it has been demonstrated that the subjective values of individual options, and not choice probabilities, are represented by neuronal firing rates.

For these two reasons we can consider choice-probability-only based systems as empirically and theoretically falsified. Of course, it may well be that groups of neurons (or the local ensemble connections of those neurons) do explicitly represent choice probabilities. Some evidence suggests that this may be the case in posterior parietal cortex, but we now have sufficient evidence to conclude that the representation of subjective values – or something much like them – occurs within the central nervous systems of primates.

Multiple Selves

The principle alternatives to the standard back-pocket model presented here are the multiple-self models that employ summation of some kind. These models typically propose the existence of two largely independent decision-making systems; one associated with so called "limbic" areas of the brain and the other with so called "rational" areas of the brain. While tremendously interesting from an economic point of view, these models are, for the most part, at variance with the majority of the existing corpus of neurobiological data . However, it is still germane to ask whether the existing evidence supports a two-agent model of decision making of the type proposed by Laibson and colleagues (see, for example, Laibson, 1997; McClure *et al*., 2004b). In that model, it is argued that the basal ganglia and medial prefrontal cortex form an emotional decision-making module which interacts (additively) with a second system organized around posterior parietal cortex and the dorsolateral prefrontal cortex, which form a rational decision-making module. Anatomical considerations that weigh against this hypothesis aside, we must ask whether or not there is compelling evidence that the division of brain areas into emotional and rational subgroups as can be supported by the available data. My answer is no. In monkeys, it has now been conclusively shown that activity in the posterior parietal cortex predicts preferences under all conditions that have been studied – for immediate rewards and for delayed rewards (Janssen and Shadlen, 2005;

Louie and Glimcher, 2006), for large rewards and for small rewards (Platt and Glimcher, 1999; Dorris and Glimcher, 2004), and for high-probability and low-probability rewards (Shadlen and Newsome, 1996; Platt and Glimcher, 1999). The data from animals seem to be unambiguous here – LIP activity predicts choices for both rational and emotional decision making. To take another example, let us turn to the basal ganglia. This is an area that a number of neuroeconomists have argued is associated with emotional decision making, but there is almost no evidence for this claim. Diseases of the basal ganglia are only very weakly associated with emotional dysfunction. The many dopaminergic forms of learning described in Part 3 of this volume, although largely mediated by the basal ganglia, do not seem to capture any clear notion of emotionality. A similar case can be made for studies of the medial prefrontal cortex. As noted above, there is evidence that this structure encodes monetary and primary rewards, preference, expected values, and gains and losses, and at least one study reports that it encodes long-delayed monetary gains. Indeed, even loss-aversion seems to be encoded in the unitary activity of this structure. Together, these data paint a picture of a structure globally involved in valuation – not a structure driven exclusively by immediacy, fear, or emotionality.

In summary, then, our available evidence seems to suggest that existing multiple-self models are largely unsupported by the bulk of our existing data. Of course, emotions do influence decision making and choosers do show varying levels of self-control; that is beyond doubt. The question is, how do emotions and circuits related to self-control affect this influence? The amygdala, to take one example, may provide an answer. The amygdala projects strongly to the ventral striatum and there is physiological and anatomical evidence that activity in the amygdala strongly influences activity in the ventral striatum. That does argue that the amygdala, and perhaps the emotions that it encodes, can influence valuation-related activity in this area, but it does not make a compelling case for a Freudian multiple-self model of neural decision making.

CONCLUSION

What emerges from a review of the available human and animal data on decision making is evidence of a two-stage model for choice. The first (or valuation) stage learns and represents the values of both actions and goods. Within this stage, at least three learning mechanisms distributed in the basal ganglia and frontal cortex contribute to the construction of what we refer to as subjective value. These areas are hypothesized to learn subjective values, at a biophysical level, through the well-studied process of synaptic plasticity. These learning processes operate both during choice and during the passive receipt of rewards, effecting a disassociation between choice and valuation.

Our available evidence makes it clear that subjective value is a stochastic quantity, effectively drawn from a spiking distribution dependent on these synaptic strengths. It is also a reference-dependent quantity, as indicated by the Tom *et al.* (2007) study. In this regard, subjective value is most closely allied to a reference-dependent random-utility model in economic theory. I propose that SV is encoded specifically in the activity of the medial prefrontal cortex and the ventral striatum. I note that while SV is responsible for preferences, it can violate the axioms of expected utility theory; indeed, it must if it is to account for true preferences. Some of these violations doubtless reflect the influence of emotion-related brain structures on medial prefrontal cortical and ventral striatal activity.

Choice, I propose, is accomplished in a network that includes the posterior parietal cortex and a number of movement-related areas subsequent to it in the motor control stream. In these areas, the SVs of objects within a single choice set are normalized to RSVs. These RSVs are further modified by the addition of a variable noise term, of Poisson-distributional origin, prior to a winner-takes-all operation that accomplishes choice itself. This is a feature reminiscent of the trembling hand of economic theory in some important ways. Let me stress that the winner-takes-all choice operation must be broadly distributed, and involves structures that range from the superior colliculus to the orbitofrontal cortex.

Of particular interest are several features of the model that remain unspecified. While there are many candidate pathways by which information from the medial prefrontal cortex and the ventral striatum may influence activity in the posterior parietal cortex, which of these pathways is critical for choice has not yet been determined. It has also been noted (see Chapter 29) that much of the posterior parietal cortex encodes SV with regard to actions, while neurons in the orbitofrontal cortex (Padoa-Schioppa and Assad, 2006), and perhaps the medial prefrontal cortex, encode SV with regard to goods. We do not know how a transformation between these representations occurs, although we do know that it does occur. We also have only limited information about the systems that "decide to choose." In some tasks animals have to be trained to make a choice as soon as possible, and under these conditions one can observe the parietal and frontal networks converging towards choice. In other situations, however, the time-courses of valuation and choice are separable, as is more typically the case in human economic behavior. This suggests the existence of a circuit that can essentially force the parietal networks towards convergence. Such a system would almost necessarily involve cortical networks of inhibitory connections, but the features of this process that decides when to choose remain completely absent from this standard back-pocket model.

Over the course of the past decade an extraordinary amount of progress has been made in identifying the basic features of the primate mechanism for choice, and there is remarkable consensus about much of this mechanism. This is a device that can be the subject of economic study, and the existing neurobiological data clearly identify some areas of economic theory as more relevant to the study of this device than others. The existing theory also identifies questions that must be answered by neurobiology. That, of course, is the whole point of this endeavor.

Acknowledgments

The author wishes to express his gratitude to Kenway Louie, Joe Kable, Ifat Levy, Daniel Burghart and Antonio Rangel for helpful discussions.

References

Allais, M. (1953). Le comportment de l'homme rationnel devant le risqué: critique des postulates et axioms de l'ecole americaine. *Econometrica* 21, 503–546.

Bayer, H.M. and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 1–13.

Breiter, H.C., Aharon, I., Kahneman, D. *et al*. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30, 619–639.

Dahlström, A. and Fuxe, K. (1964). Evidence for the existence of monoamine-containing neurons in the central nervous system. I. Demonstration of monoamines in the cell bodies of brain stem neurones. *Acta Physiol. Scand.* 62(Suppl. 1), 1–55.

Daw, N.D., Kakade, S., and Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Networks* 15, 603–616.

Delgado, M.R., Nystrom, L.E., Fissell, C. *et al*. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *J. Neurophysiol.* 84, 3072–3077.

Dorris, M.C. and Glimcher, P.W. (2004). Activity in posterior parietal cortex is correlated with the subjective desirability of an action. *Neuron* 44, 365–378.

Elliott, R., Friston, K.J., and Dolan, R.J. (2000). Disasociable neural responses in human reward systems. *J. Neurosci* 20, 6159–6165.

Elliott, R., Newman, J.L., Longe, O.A., and Deakin, J.F.W. (2003). Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. *J. Neurosci.* 23, 303–307.

Glimcher, P.W. (2003). The neurobiology of visual saccadic decision making. *Annu. Rev. Neurosci.* 26, 133–179.

Glimcher, P.W. (2005). Indeterminacy in brain and behavior. *Annu. Rev. Psychol.* 56, 25–56.

Glimcher, P.W. and Sparks, D.L. (1992). Movement selection in advance of action in the superior colliculus. *Nature* 355, 542–545.

Glimcher, P.W., Kable, J.W., and Louie, K. (2007). Neuroeconomic studies of impulsivity: now or just as soon as possible? *Am. Econ. Rev.* 97, 142–147.

Gold, J.I. and Shadlen, M.N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature* 404, 390–394.

Gul, F. and Pesendorfer, W. (2008). The case for mindless economics. In: A. Caplin and A. Schotter (eds), *The Foundations of Positive and Normative Economics: A Handbook*. Oxford: Oxford University Press, (forthcoming).

Heeger, D.J. (1992). Normalization of cell responses in cat striate cortex. *Vis. Neurosci.* 9, 81–198.

Hollerman, J.R. and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309.

Iyengar, S. and Lepper, M. (2000). When choice is demotivating: can one desire too much of a good thing? *J. Pers. Social Psych.* 79, 995–1006.

Iyengar, S.S., Wells, R.E., and Schwartz, B. (2006). Doing better but feeling worse: looking for the "best" job undermines satisfaction. *Psychol. Sci.* 17, 143–150.

Janssen, P. and Shadlen, M.N. (2005). A representation of the hazard rate of elapsed time in macaque area LIP. *Nat. Neurosci.* 8, 234–241.

Kable, J.W. and Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291.

Knutson, B., Westdorp, A., Kaiser, E., and Hommer, D. (2000). fMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage* 12, 20–27.

Knutson, B., Fong, G.W., Bennett, S.M. *et al*. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related FMRI. *NeuroImage* 18, 263–272.

Knutson, B., Taylor, J., Kaufman, M. *et al*. (2005). Distributed neural representation of expected value. *J. Neurosci.* 25, 4806–4812.

Krug, K., Cumming, B.G., and Parker, A.J. (2004). Comparing perceptual signals of V5/MT neurons in two binocular depth tasks. *J. Neurophysiol.* 92, 1586–1596.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Q. J. Economics* May, 443–477.

Lau, B. and Glimcher, P.W. (2008). Value representations in the primate striatum during matching behavior. *Neuron* 58, 451–643.

Lee, C., Rohrer, W.H., and Sparks, D.L. (1989). Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature* 332, 357–360.

Levy, I., Schluppeck, D., Heeger, D.J., and Glimcher, P.W. (2007). Specificity of human cortical areas for reaches and saccades. *J. Neurosci.* 27, 4687–4696.

Louie, K. and Glimcher, P.W. (2006). Temporal discounting activity in monkey parietal neurons during intertemporal choice. *Soc. Neurosci. Abstr.*, 605.5.

Louie, K., Grattan, L., and Glimcher, P.W. (2007). Relative reward encoding and cortical normalization in parietal area LIP. *Soc. Neurosci. Abstr.*, 645.7.

McClure, S.M., Laibson, D.I., Loewenstein, G., and Cohen, J.D. (2004a). Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507.

McClure, S.M., Li, J., Tomlin, D. *et al*. (2004b). Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* 44, 379–387.

McFadden, D. (1974). Conditional Logit analysis of qualitative choice behavior. In: P. Zarembka (ed.), *Frontiers in Econometrics*. New York, NY: Academic Press, pp. 105–142.

Montague, P.R., Dayan, P., and Sejnowski, T.J. (1997). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.

Newsome, W.T., Britten, K.H., and Movshon, J.A. (1989). Neuronal correlates of a perceptual decision. *Nature* 341, 52–54.

O'Doherty, J., Deichmann, R., Critchley, H.D., and Dolan, R.J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron* 33, 815–826.

O'Doherty, J.P., Buchanan, T.W., Seymour, B., and Dolan, R.J. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron* 49, 157–166.

Padoa-Schioppa, C. and Assad, J.A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.

Padoa-Schioppa, C. and Assad, J.A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes in menu. *Nat. Neurosci.* 11, 95–102.

Pagnoni, G., Zink, C.F., Montague, P.R., and Berns, G.S. (2002). Activity in human ventral striatum locked to errors in reward prediction. *Nat. Neurosci.* 5, 97–98.

Plassmann, H., O'Doherty, J., and Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27, 9984–9988.

Platt, M.L. and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.

Platt, M.L., Lau, B., and Glimcher, P.W. (2003). Situating the superior colliculus within the gaze control network. In: W.C. Hall and A. Moschovakis (eds), *The Oculomotor System: New Approaches for Studying Sensorimotor Integration*. Boca Raton, FL: CRC Press, pp. 1–34.

Romo, R. and Schultz, W. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. *J. Neurophysiol.* 63, 592–606.

Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340.

Savage, L. (1954). *Foundations of Statistics*. New York, NY: Wiley.

Schultz, W. and Romo, R. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioral reactions. *J. Neurophysiol.* 63, 607–624.

Schultz, W., Apicella, P., and Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.* 13, 900–913.

Schwartz, O. and Simoncelli, E.P. (2001). Natural signal statistics and sensory gain control. *Nat. Neurosci.* 4, 819–825.

Selten, R. (1975). A reexamination of the perfectness concept for equilibrium points in extensive games. *Intl J. Game Theory* 4, 25–55.

Shadlen, M.N. and Newsome, W.T. (1996). Motion perception: seeing and deciding. *Proc. Natl Acad. Sci. USA* 93, 628–633.

Stevens, S. (1970). Neural events and the psychophysical law. *Science* 170, 1043–1050.

Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.

Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642–1645.

Tolhurst, D.J., Movshon, J.A., and Dean, A.F. (1983). Statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res.* 23, 775–785.

Tom, S.M., Fox, C.R., Trepel, C., and Poldrack, R.A. (2007). The neural basis of loss aversion in decision-making under risk. *Science* 315, 515–518.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

CHAPTER

33

Remarks on Neuroeconomics

Daniel Kahneman

I have been an enthusiastic and ignorant fan, since the early days, of what used to be called cognitive neuroscience, and an even more ardent fan since an event that happened, so it seems, yesterday – the skyrocketing takeoff of neuroeconomics. Over the years I have had many difficult conversations with colleagues in my discipline, who share my ignorance but not my enthusiasm. The skeptics are quick to agree that thinking and choice are outcomes of events that occur in the brain; they question the usefulness *to psychology* of knowing which parts of the brain are activated when particular thoughts or choices happen[1].

Imaging studies provide the main source of data for neuroeconomics, although there is hope that this could change as new techniques are introduced. To be useful to psychology, of course, such measures of brain activity must be translated into psychological terms. One of the happy surprises of neuroeconomics is the frequent finding of impressive correlations between psychological measures and measures of brain activity (or, sometimes, of differences between levels of activity in different regions). Among other examples cited in this book are a within-subject correlation between the rejection of unfair offers and the activation of right anterior insula (Sanfey and Dorris), a high correlation between a behavioral measure of loss aversion and a neural measure of loss aversion in ventral striatum and ventrolateral prefrontal cortex (Fox and Poldrack), and a correlation between the *difference* in neural measures of people's utility for their own money and for others' money and a willingness to donate to these others (see Chapter 20 of this volume). The within-subject correlations suggest an impressive ability to identify a close correlate of a causal mechanism.

The between-subject correlations help to make another point, on a topic on which psychologists and economists have very different perspectives: the question of interpersonal comparisons of utility. If brains are sufficiently similar in structure and function to support high correlations between indices of brain activity and measures of psychological states, the calibration function that relates the psychological to the neural variables is unlikely to be very different across individuals, and there appears to be little justification for the taboo regarding interpersonal comparisons.

High correlations between well-identified psychological and neural measures are the exception, not the norm. In most experiments, of course, the correspondence between psychological terms and neural measures is more equivocal, and the interpretation of imaging results is tricky. Russ Poldrack (see Chapter 11 of this volume; also Poldrack, 2006) has drawn attention to the problem of "reverse inference," which arises when people infer a specific psychological process from activity in a particular region – for example,

---

[1] The argument is strikingly similar to the case that some economists have made in favor of a "mindless economics": psychological data should be ignored because economics is a self-contained discipline (Gul and Pesendorfer, 2005). A case for "brainless psychology" is even easier to make.

when activity in dorsal striatum is interpreted as an indication that people enjoy punishing strangers who have behaved unfairly (De Quervain *et al.*, 2004; see also Chapter 17 of this volume). There is indeed a problem, because activity in ventral striatum is not perfectly correlated with enjoyment: many other circumstances produce activity in that region, and there is no assurance that it will be active whenever the individual experiences pleasure. In spite of this difficulty, the result and its proposed interpretation are just what a general psychologist (not a neuroscience specialist) would order. It is surprising but plausible, and it drives thinking in new directions. The more difficult test, for a general psychologist, is to remember that the new idea is still a hypothesis which has passed only a rather low standard of proof. I know the test is difficult, because I fail it: I believe the interpretation, and do not label it with an asterisk when I think about it. And I will be sorry if it is disproved, but will have no difficulty in accepting its demise – it would join a long list of defunct once-cherished ideas.

In spite of the fact that many regions of the brain respond promiscuously to many experimental situations, the interpretation of positive results seems relatively tractable. I like the analogy of a personality inventory, a collection of hundreds of questions with dichotomous answers (my memory from graduate school is that the item "I used to enjoy ripping the wings off flies" (or words to that effect), when answered in the negative is a weak indication of depression). Every brain region is like an individual, answering "yes" or "no" to thousands of experimental questions that are devised by the scientific community. The *pattern* of yes–no answers – where significant activation is "yes" – is an indication of the "personality" of a region. As the evidence accumulates, it seems very likely that practitioners of the art will develop an intuitive appreciation of these personalities – a sense of how a region is likely to respond to a new situation – which should help in the critical stage of selecting regions of interest. It is also likely that computational techniques will be applied in an attempt to organize the explosively growing amount of positive findings and ideas.

The interpretation of negative results presents a more difficult challenge. The asymmetry between the treatment of positive and negative findings is a major feature of the current situation, probably unavoidable in the early days of a developing methodology. However, it is a problem, as I well remember from my only foray into the field (Breiter *et al.*, 2001). Loss aversion is a salient aspect of the behavioral analysis of decisions, and I had expected to see manifestations of loss aversion showing up in vivid colors on the images. I had also anticipated that outcomes would be more salient than expectations – which they are, of course, in our subjective experience of life. None of this happened. Gains showed up more than losses, and the sight of a favorable or unfavorable prospect caused more activation than its eventual resolution. These were clearly important observations, but I did not know what we could make of them that would change psychological thinking about decision making. The similarity to the reactions of the dopamine system (see Chapter 21) provided a context for the relative salience of predictions, although it did not resolve the apparent inconsistency with subjective experience. The absence of obvious manifestations of the powerful psychological response to losses has remained a problem, although several chapters in the book report advances (see Chapters 7, 11, and 12; see also Knutson *et al.*, 2007).

Early in the days of any field or any technique, the most valuable results are those that confirm expectations or help to sharpen existing hypotheses. As the field and its techniques mature, of course, believable surprises become the goal and the prize. My impression is that, at least in the domains with which I am familiar, it is still early days for neuroeconomics. When I was asked by a well-known critic of the field what I had learned that had changed my mind about decision making, I did not have much to say. In part, of course, it is because I was right all along! The findings of neuroeconomics research have generally confirmed the expectations of behavioral decision theorists and behavioral economics. However, we are beginning to learn more, and I am confident that the pace will accelerate in coming years. One example of what we are learning concerns the interpretation of framing effects. When people are presented with the same choice between a sure thing and a gamble, they prefer the sure thing if the outcomes are framed as gains and they prefer the gamble if the outcomes are phrased as losses. Why? Some recent findings and theoretical analyses suggest an interpretation: a primary tendency to hate sure losses and to be attracted to sure gains, when such simple outcomes are compared to more complicated risky prospects. In this volume, the idea that these preferences are due to a simple Pavlovian response is suggested by both Dayan and Seymour (Chapter 13), and Bossaerts, Preuschoff, and Hsu (Chapter 23). An imaging study of framing effects (De Martino *et al.*, 2006) led Shane Frederick and myself to a similar interpretation (Kahneman and Frederick, 2007), which gains further indirect support from the ingenious studies of framing effects in capuchin monkeys that Santos and Chen report in Chapter 7 of this volume. Another topic on which neural data may help to articulate theory concerns the rewards that guide actions. This is one of the main issues of contention between behavioral economics and the standard economic model. The notion that preferences are constructed in the immediate context stands in sharp contrast to the idea of a preference order which is basic to rational-choice theories. The concept of actions that are their own rewards – altruistic punishment and charitable giving may be examples – is, of course, much more compatible with the former than with the latter, and there is growing support for it in neural data (see Chapters 6, 15, and 20). The very meaning of rationality changes when actions are determined by the intrinsic utility they provide. I believe we are well on our way to the day where our theoretical concepts about decision making are shaped at least in part by findings from neuroscience.

The difficulties associated with the conceptual interpretation of neuroscience evidence can be traced to the correlational nature of most of that evidence. A new era in neuroeconomics began with the introduction of experimental manipulations such as transcranial direct current stimulation (tDCS) or the inhalation of oxytocin to induce predictable behavioral consequences (see Chapter 15). Fehr's justifiable enthusiasm is captured in this quotation:

> A key feature of tDCS is that it is inexpensive and can be simultaneously applied to many subjects who interact in a laboratory environment (Knoch *et al.*, 2007). Thus, in principle, tDCS can be applied to a group of, say, 20 subjects simultaneously, with each of them playing one one-shot game with the other 19 subjects. Therefore, tDCS could prove to be a non-invasive brain stimulation method that revolutionizes neuroeconomics because it greatly enhances data-collection efficiency and enables brain stimulations in whole groups of interacting subjects.

The application of minimally invasive interventions that alter brain function will drastically increase the relevance of studies of the brain to an understanding of the deciding mind. The only development that could do as much for the field would be a qualitative improvement in the temporal resolution for patterns of activation.

The instructive chapter that introduced this volume mentioned two studies published in 2001 (Breiter *et al.*, 2001; McCabe *et al.*, 2001) that announced (they did not cause) two strands of research which are prominent in this collection: studies of social preference in contexts defined by game theory, and studies of valuation of uncertain outcomes. To say that a great deal has happened since then is an obvious understatement, but the progress has not been even. The valuation and decision problem is broader and harder, and a meaningful theoretical integration seems much more remote. For one thing, there is no agreement on whether the theoretical language for such an integration will come from computational neuroscience, from rational choice theory, or from behavioral economics. In Chapter 32, Paul Glimcher offers utility theory as a guiding framework for the field, while acknowledging that there is much it does not cover. His proposal will certainly be contested. A unitary concept of utility cannot do justice to the complexities of the relationships between what people (and other animals) want, what they expect to enjoy, what they actually enjoy, and what they remember having enjoyed. Unless one believes that there is never a difference in these types of utility for anybody's decision, neuroeconomics might provide some of the best tools to learn when they differ and why. Furthermore, the hypothesis of optimal performance is not equally applicable to all domains. We know, for example, that the human perceptual system is more reliably Bayesian than is human judgment, and optimality in the organization of movement (see Chapter 8) or in the automatic allocation of attention does not necessarily generalize to other categories of performance.

What do we do when behavioral evidence and neural evidence point in different directions? Some time in the future, I expect, such discrepancies may come to be resolved by reinterpreting the behavioral data to fit the conclusions from the neuroscientific work. This day has not come yet. Here again, in Chapter 32, Glimcher raises an interesting challenge. He believes that there is no evidence from neuroscience to support the idea that (some) decisions arise from a conflict between emotion and reason, and concludes that a unitary system will do the job. His position on the neuroscience is not universally held (see, for example, McClure *et al.*, 2007), but this is not my point. I start from the position that there is overwhelming behavioral evidence both for the existence of multiple systems of thought and choice and for the importance of conflict (Kahneman, 2003). As of today, the absence of repeated studies showing well-localized neuroscientific correlates for multiple systems should still be viewed as an unsolved problem for neuroeconomics, rather than as a problem for the idea of multiple systems. I emphasize "as of today," because I firmly believe that the findings of neuroscience – negative as well as positive – will soon play a large role in shaping the concepts and theories of behavioral research.

Acknowledgments

I thank Colin Camerer, Craig Fox and Russ Poldrack for helpful comments and suggestions. The usual caveats apply.

References

Breiter, H.C., Aharon, I., Kahneman, D. *et al*. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30, 619–639.

De Martino, B., Kumaran, D., Seymour, B., and Dolan, R.J. (2006). Frames, biases, and rational decision-making in the human brain. *Science* 313, 684–687.

De Quervain, D.J.-F., Fischbacher, U., Treyer, V. *et al*. (2004). The neural basis of altruistic punishment. *Science* 305, 1254–1258.

Kahneman, D. (2003). Maps of bounded rationality: psychology for behavioral economics. *Am. Econ. Rev.* 93, 1449–1475.

Kahneman, D. and Frederick, S. (2007). Frames and brains: elicitation and control of response tendencies. *Trends Cogn. Sci.* 11, 45–46.

Knoch, D., Nitsche, M.A., Fischbacher, U. *et al*. (2007). Studying the neurobiology of social interaction with transcranial direct current stimulation: the example of punishing unfairness. *Cerebral Cortex Advance Access*.

Knutson, B., Rick, S., Wirnmer, G.E. *et al*. (2007). Neural predictors of purchases. *Neuron* 53, 147–156.

McCabe, K., Houser, D., Ryan, L. *et al*. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl Acad. Sci. USA* 98, 11832–11835.

Poldrack, R.A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10, 59–63.