PRINCIPLES OF ADDICTION

Comprehensive Addictive Behaviors and Disorders

Volume 1

Editor-in-Chief

PETER M. MILLER
Psychiatry and Behavioral Sciences,
Medical University of South Carolina,
Charleston, SC, USA

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Academic Press is an imprint of Elsevier

Academic Press
Academic Press is an imprint of Elsevier
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA
32 Jamestown Road, London NW1 7BY, UK
225 Wyman Street, Waltham, MA 02451, USA

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

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

ISBN: 978-0-12-398336-7

Printed and bound in the United States of America
13 14 15 16 17 10 9 8 7 6 5 4 3 2 1

Preface

Principles of Addiction is one of three volumes encompassing the 2500 page series.

Comprehensive addictive behaviors and disorders: In both print and online formats, this series provides the most comprehensive compilation of current knowledge on addictive behaviors and disorders to date. In short, it is the definitive reference work on addictions.

The significance of this series stems from the fact that addictive behaviors and disorders represent major personal, social, and public health problems throughout the world. While research on addictions has grown exponentially over the past 20 years, the primary literature in this field is widely dispersed. For researchers and clinicians, staying abreast of this vast and expanding knowledge is a challenging, if not impossible, task. Researchers specializing in one addiction subspecialty (e.g. clinical research, neuroscience, health services, public policy, treatment, pharmacology, genetics) are unable to keep apprised of the big picture. In addition, research findings on one type of addiction (e.g. alcohol dependence, excessive gambling, methamphetamine dependence) may have relevance to other types (e.g. cocaine or opiate dependence) but such findings are not readily available to all addiction scientists since they are dispersed among so many specialty journals, books and web sites. Until now, there has been no all-encompassing resource that could serve as the "go-to" compendium for information on any and all addictions.

Comprehensive addictive behaviors and disorders fills this void by providing a unique and valuable storehouse of interdisciplinary scientific information for researchers, clinicians, and policy makers that comprehensively summarizes state-of-the-art knowledge. The terms "behavior" as well as "disorder" are used purposefully in the title since the work includes both excessive use (what is often called "at risk" or "harmful" use) as well as true physiological dependence. In addition, there is lack of evidence and some controversy over the issue of whether or not some behavioral excesses (e.g. texting or video gaming) can be classified as true addictions. Finally a number of the more basic, translational chapters cover addictive brain or behavioral processes rather than focusing on "disorders" per se.

Conceptually, the three volumes that make up this series, *Principles of Addiction*, *Biological Research on Addiction*, and *Interventions for Addictions*, cover an extensive range of topics including, but not limited to, the nature of addiction, cravings, comorbidities, types of addictions, behavioral biology, neuroscience, neuroimaging, genetics, neuropharmacology, psychosocial treatments, addiction medications, application of addiction science to practice, public policy, and prevention. With the growing emphasis on translational research, the goal has been to integrate diverse findings into a meaningful conceptualization of all aspects of use and abuse.

The audience for the series includes advanced undergraduates, graduates and postdoctoral students, professors, researchers, clinicians and policy makers. The series can also serve as a valuable aid to instructors and students in the hundreds of university-level addiction degree programs throughout the world. As an encyclopedic series, the mass media as well as the general public will find this work to be a comprehensive source of evidence-based, scientific information on addictions. This is especially important since the field of addiction is continually plagued by anecdotal and misleading information found both in print and online.

SCOPE AND FORMAT

The development and compilation of this series have been a truly collaborative effort. Nine internationally recognized addiction experts have served on the editorial board for this project. The board is composed of Peter Miller, PhD, Medical University of South Carolina; Mary Larimer, PhD, University of Washington; Kyle Kampman, MD, University of Pennsylvania School of Medicine; David Kavanagh, PhD, Queensland University of Technology; Samuel Ball, PhD, Yale University School of Medicine; Phillipe DeWitte, MD, Université Catholique de Louvain (Belgium); Marsha Bates, PhD, Rutgers University; Nancy Petry, PhD, University of Connecticut Health Center; and Arthur Blume, PhD, Washington State University. The board members were chosen to provide a wide range of interdisciplinary expertise.

The board members were responsible for selecting and inviting authors for chapters in their specialty areas and to provide editorial guidance. Chapters are authored by reputable, well-recognized authorities in the addictions field. Authors were selected for their

expertise and experience, with particular emphasis placed on selecting an international group with diverse philosophies and research backgrounds.

Each chapter is approximately 6500 words in length. Chapters consist of approximately 10 printed pages each, ranging from 5 to 15 pages depending on the topic. Many chapters include tables and figures to better illustrate data. Authors were asked to summarize current knowledge in their areas without providing references within the text. Thus, the work reads like an encyclopedia, providing the reader with an overview of the state-of-the-art rather than an in-depth research report. A further reading and web site list are provided at the end of each chapter for those who require research references and more detailed information. In addition, an outline, glossary, list of keywords, and list of cross-references are provided for each chapter.

The hope is that this compendium will provide a universal platform for a more science-based approach to the study, prevention and treatment of addictions. The ultimate goal is to improve the lives of addicted individuals and their families throughout the world through a more comprehensive and detailed understanding of the addictive process.

VOLUME 1: PRINCIPLES OF ADDICTION

This first volume, *Principles of Addiction*, provides scientific information in two broad categories: the nature of addiction and types of addiction. The volume sets off the series by giving an overview of what is known about the addictive process as well as a depiction of each of the known addictive behaviors and disorders.

Chapters related to the nature of addiction are designed to provide basic descriptions of current knowledge of the addictive process in terms of its initial development, mechanisms of action, and basic characteristics. Chapters cover the history of the study and treatment of addiction, its epidemiology as well as its basic definitions, diagnostic criteria, and theoretical models. Essential characteristics of dependence, including loss of control, tolerance, craving, withdrawal, denial, and relapse, are explained and discussed. Chapters cover a wide range of information regarding etiology and developmental course. The influence of gender, stress, family, and peer influence is explained. Associated features such as stress and psychiatric comorbidities (e.g. depression) are included. Coverage of many of these issues includes both excessive use of substances as well as true dependence and why some people "mature out" of these behaviors and others become entrapped in severe addiction.

To demonstrate the all-inclusive nature of this volume, there are chapters on 27 different addictions, both substance-focused and behavioral. Topics of substance-focused use and abuse include alcohol, heroin, cocaine marijuana, hallucinogens, ecstasy/MDMA, inhalants, ketamine, anabolic-androgenic steroids, prescription and over-the-counter medications, tobacco, caffeine and energy drinks, food, areca nut, khat, and water pipe smoking. Behavioral excesses include gambling, shopping, exercise, work, and sex as well as behaviors that have become potential problems in contemporary society such as cybersex, virtual worlds on the internet, video games, cell phone use and texting, and social networking. For both substance-related and behavioral excesses authors were asked to address the issue of whether or not there is enough scientific evidence to classify each as a true "addiction." For example, there is controversy over whether or not excessive sexual activity and food consumption constitute true addictions from a diagnostic sense. Unfortunately, scientific study of many of the behavioral addictions is so recent (with the exception of gambling) that little is known about these syndromes.

A few substances are so new that little is known about them. They were not included in this volume since no major human studies have yet been conducted on their use, mechanisms of action, abuse potential, or health effects. For example, the National Institute on Drug Abuse (NIDA) describes three such emerging drugs: "spice," an herbal mixture with effects similar to marijuana; salvia, an herb that causes hallucinations and psychotic-like symptoms; and "bath salts," a synthetic compound with amphetamine-like effects that result in increased blood pressure and heart rate. In time, more scientific knowledge will be accumulated on these substances but, at this stage, it is apparent that the drug culture is able to stay one step ahead of the addiction specialists in developing new compounds to use and abuse.

In summary, this volume provides a basic understanding of the definitional and diagnostic differences between use, abuse, and disorder. It describes in great detail the characteristics of these syndromes and various etiological models. Finally, it provides detailed descriptions of a variety of addictive behaviors and disorders and their similarities and differences. This volume is especially important in providing a basic introduction to the field as well as an in-depth review of our current understanding of the nature and process of addictive behaviors.

Peter M. Miller

Editors: Biographies

Peter Miller

Peter M. Miller, PhD, is a professor of psychiatry and behavioral sciences in the Center for Drug and Alcohol Programs at the Medical University of South Carolina. He also holds a faculty appointment in the College of Dental Medicine. He is a clinical psychologist and is board certified by the American Board of Professional Psychology. He specializes in research on alcohol and substance abuse screening and intervention in medical and dental settings, with particular reference to alcohol-sensitive diseases such as hypertension and oral cancer. He has published over 100 scientific articles, has authored 11 books and has served as editor on major addictions textbooks and reference works. He is editor in chief of two international research journals, *Addictive Behaviors* and *Eating Behaviors*, and serves on several editorial boards. He is past president of the International Society of Addiction Journal Editors (ISAJE).

Samuel Ball

Samuel A. Ball, PhD, is a professor and assistant chair for education and career development in the Department of Psychiatry at Yale University School of Medicine. He also serves as research director for the NIDA-funded Psychotherapy Development Research Center, NIH-funded BIRCWH Women's Health and Addictive Behaviors program, and The APT Foundation in New Haven, CT. His research focuses on the assessment and treatment implications of personality dimensions, personality disorders, and multidimensional subtypes in substance abuse.

Arthur Blume

Arthur W. Blume is a professor of psychology at Washington State University. Before joining the faculty at Washington State University, Dr Blume was on the faculty at the University of North Carolina at Charlotte and before that at the University of Texas at El Paso. He currently serves as a section editor for this volume, and as an associate editor of the journals *Addictive Behaviors* and *Cultural Diversity and Ethnic Minority Psychology*. From 2007 to 2009, he was an American Indian representative on the National Committee on Ethnic Minority Affairs of the American Psychological Association in the United States. His program of research has focused on addictive behaviors among high-risk populations, especially those from ethnic minority groups in the United States. His extramurally funded research and publications reflect his keen interests in ethnic minority health and well-being.

David Kavanagh

David Kavanagh, PhD, is a research capacity-building professor in the Institute and Health & Biomedical Innovation and School of Psychology & Counselling at Queensland University of Technology, and has adjunct posts at the University of Queensland and Griffith University. He was educated at Sydney and Stanford Universities and led a community mental health service before becoming an academic. He has researched and written widely on addiction, comorbidity and dissemination of evidence-based treatments and is currently researching the elicitation and maintenance of functional motivation. He has been on the editorial boards of several journals, including *Addiction and Addictive Behaviors*, and has served on several state and national expert committees on comorbidity and addiction.

Kyle Kampman

A board-certified psychiatrist specializing in addiction psychiatry, Dr Kyle M. Kampman, MD, received his medical degree from Tulane University School of Medicine. He completed his residency in psychiatry and fellowship in addiction psychiatry at the University of Pennsylvania. Currently, he is a medical director of the Charles O'Brien Center for the Treatment of Addictions, professor of psychiatry, and medical director of the Treatment Research Center at the University of Pennsylvania in Philadelphia. His research interests include pharmacotherapy for cocaine dependence and the cocaine withdrawal syndrome. He is an associate editor of Drug and Alcohol Dependence. He serves on the editorial board of the *Journal of Addiction Medicine* and is an active member of the American Society of Addiction Medicine, College on Problems of Drug Dependence, and the Pennsylvania Society of Addiction Medicine. His contributions to the field have been acknowledged by the Scott Mackler Award for Excellence in Substance Abuse Teaching presented by the University of Pennsylvania School of Medicine and the Caron Foundation's Medical Professional-Physician Award.

Marsha E. Bates

Marsha E. Bates, PhD (Rutgers – the State University of New Jersey), is a research professor of psychology at the Center of Alcohol Studies (CAS) at Rutgers

University and an associate professor of psychiatry at UMDNJ/Robert Wood Johnson Medical School. She directs the Cognitive Neuroscience Laboratory that promotes translation between basic human experimental and clinical science. Her current multiinstitution research project is a component of the NIAAA Mechanism of Behavior Change Interdisciplinary Research Consortium (MIRC) which seeks to build novel approaches to alcohol-related problems via mechanism-based strategies. Dr Bates is vice-chair of the Rutgers Institutional Review Board for the Protection of Human Subjects involved in research. She is a member of the Behavioral and Social Advisory Council of ABMRF/The Foundation for Alcohol Research and vice-chair of the Board of Trustees of Alcohol Research Documentation, Inc. She is an editorial board member of the *Journal of Studies on Alcohol and Drugs* and a past associate editor of Psychology of Addictive Behaviors and Alcoholism: Clinical and Experimental Research. She is a fellow of the American Psychological Association (APA), previously served as president of APA Division 50 (Society of Addiction Psychologists), and received their 2011 Distinguished Scientific Contribution Award. Address: Center of Alcohol Studies, Rutgers – the State University of New Jersey, 607 Allison Road, Piscataway, NJ 08854, USA; E-mail: mebates@rutgers.edu.

Mary Larimer

Mary E. Larimer is a professor of psychiatry and behavioral sciences, an adjunct professor of psychology, associate director of the Addictive Behaviors Research Center, and director of the Center for the Study of Health and Risk Behaviors at the University of Washington. She received her PhD in clinical psychology from the University of Washington and has been a member of the faculty since 1995. Dr Larimer's research and clinical interests include (1) prevention and treatment of alcohol and drug problems among adolescents and young adults (with a particular focus on college drinking prevention), (2) cross-cultural research regarding prediction of initiation of drinking and trajectories of alcohol and substance use during emerging adulthood, (3) comorbidity of substance use with depression, suicide, trauma, PTSD, disordered eating, and gambling problems, (4) evaluation of housing and treatment programs for chronically homeless and incarcerated individuals, and (5) dissemination of evidence-based prevention and treatment approaches into clinical-, school-, and work-site settings. She has published more than 100 articles and book chapters on these topics.

Nancy M. Petry

Nancy Petry is a professor of medicine at the University of Connecticut Health Center, and she earned her PhD in psychology from Harvard University. She developed the prize reinforcement system that has been widely disseminated in the context of treating substance use disorders and is now being applied to address other conditions including overweight/obesity, exercise, medication adherence, and diabetes. She also conducts research related to pathological gambling and its treatment.

Philippe De Witte

Professor De Witte heads the Laboratory of Behavioural Biology at the Université Catholique de Louvain. In 1987, he received a Fulbright grant to complete a specialization in brain research at the NIH. He has also worked as invited professor in Washington State University and the University of Colorado.

He is the editor in chief of *Alcohol and Alcoholism* and on the editorial advisory board of a number of leading journals. He was president of the European Society for Biomedical Research on Alcoholism (ESBRA) for two terms from 1993 to 2001 and was the president of the International Society for Biomedical Research on Alcoholism (ISBRA, 1998–2002). He joined the ERAB Advisory Board in 2003 and became chairman on 1 January 2007.

His research interests include pharmacology and neurobiology of addiction including therapeutics and treatments. He is a member of a number of professional bodies including ESBRA, ISBRA, RSA, The Belgian College of Neuropsychopharmacology and Biological Psychiatry (BCNBP) and the Société française d'Alcoologie (SFA) and has published over 200 articles in scientific journals, twice as many abstracts and book chapters.

List of Contributors

Jeremy Adams Eclectic Consulting Pty. Ltd., Altona, VIC, Australia

Christina Aivadyan New York State Psychiatric Institute, NYSPI/Columbia University, New York, NY, USA

Jackie Andrade University of Plymouth, Drake Circus, Plymouth, UK

Cecilie Schou Andreassen The Bergen Clinics Foundation, Norway , Department of Psychosocial Science, University of Bergen, Norway

Angela Attwood School of Experimental Psychology, University of Bristol, Bristol, UK

Paul Aveyard Primary Care Clinical Sciences, University of Birmingham, Edgbaston, UK

Amanda Baker Center for Brain and Mental Health Research, University of Newcastle, Callaghan, NSW, Australia

Kimberly F. Balsam Palo Alto University, Palo Alto, CA, USA

John Bancroft The Kinsey Institute for Research in Sex, Gender and Reproduction, Indiana University, Bloomington, IN, USA

Kelly S. Barth Medical University of South Carolina, Charleston, SC, USA

Linda Bauld School of Management, University of Stirling, Stirling, UK

Marta Beranuy FPCEE Blanquerna, Universidad Ramon Llull, Barcelona, Spain

Arthur W. Blume Washington State University, Vancouver, WA, USA

James M. Bolton PsycHealth Centre, Winnipeg, MB, Canada

Marcel O. Bonn-Miller National Center for PTSD and Center for Health Care Evaluation, VA Palo Alto Health Care System, Palo Alto, CA, USA

Brian Borsari Brown University, Providence, RI, USA

Kathleen T. Brady Medical University of South Carolina, Charleston, SC, USA

Ashley R. Braun University of Illinois at Chicago, Chicago, IL, USA

John Britton Division of Epidemiology and Public Health, Clinical Sciences Building, Nottingham City Hospital, Nottingham, UK

Brian P. Brown University of North Carolina Wilmington, Wilmington, NC, USA

Justine Campbell University of Queensland, St. Lucia, QLD, Australia

Xavier Carbonell FPCEE Blanquerna, Universidad Ramon Llull, Barcelona, Spain

Félix Carvalho University of Porto, Porto, Portugal

Gabriele Caselli Studi Cognitivi, Modena, Italy

Natalie Castellanos-Ryan Université de Montréal, Montreal, QC, Canada

Grace Chan University of Connecticut School of Medicine, Farmington, CT, USA

Karen Grube Chartier University of Texas, Dallas, TX, USA

Chuan-Yu Chen Institute of Public Health, National Yang-Ming University, Taiwan; Division of Mental Health and Addiction Medicine Research, National Health Research Institutes, Taiwan

Wei J. Chen Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taiwan

Tammy Chung University of Pittsburgh Medical Center, Pittsburgh, PA, USA

Kelly J. Clemens Macquarie University, Sydney, NSW, Australia

Anahi Collado-Rodriguez University of Maryland College Park, College Park, MD, USA

Megan Conrad University of Illinois at Chicago, Chicago, IL, USA

Patricia J. Conrod Université de Montréal, Montreal, QC, Canada

Jan Copeland National Cannabis Prevention and Information Centre, University of New South Wales, Sydney, NSW, Australia

Christopher J. Correia Auburn University, Auburn, AL, USA

Baine B. Craft Seattle Pacific University, Seattle, WA, USA

Ann Crosland Sunderland University, Sunderland, UK

Teresa Cunha-Oliveira CNC, University of Coimbra, Coimbra, Portugal

Kelly C. Davis University of Washington, Seattle, WA, USA

Mark Deady National Drug and Alcohol Research Centre, University of New South Wales, Sydney, NSW, Australia

Louisa Degenhardt University of New South Wales, Sydney NSW, Australia

Paul H. Delfabbro The University of Adelaide, Adelaide, SA, Australia

Ashley A. Dennhardt University of Memphis, Memphis, TN, USA

Paul Dillon National Cannabis Prevention and Information Centre, University of New South Wales, Sydney, NSW, Australia

David J. Drobes Moffitt Cancer Center, University of South Florida, Tampa, FL, USA

Jessica L. Eaddy Clinical Neuroscience Division, Medical University of South Carolina, Charleston, SC, USA

Enrique Echeburúa University of the Basque Country UPV/EHU, Avda de Tolosa, 70 San Sebastián, Spain

Vivian B. Faden National Institute on Alcohol Abuse and Alcoholism, Bethesda, MD, USA

Miriam C. Fenton Columbia University, Mailman School of Public Health, New York State Psychiatric Institute, NYSPI/Columbia University, New York, NY, USA

Stuart G. Ferguson School of Pharmacy, University of Tasmania, Hobart, TAS, Australia; Menzies Research Institute Tasmania, University of Tasmania, Hobart, TAS, Australia

M.J. Fernandez-Serrano Department of Psychology, Universidad de Jaén, Jaén, Spain

Matt Field University of Liverpool, Liverpool, UK

Nicole Fossos University of Washington, Seattle, WA, USA

Dawn W. Foster University of Houston, Houston, TX, USA

Ingmar H.A. Franken Institute of Psychology, Erasmus University Rotterdam, Rotterdam, The Netherlands

Irene M. Geisner Center for the Study of Health and Risk Behaviors, University of Washington, Seattle, WA, USA

William H. George University of Washington, Seattle, WA, USA

Lilian A. Ghandour American University of Beirut, Riad El-Solh, Beirut, Lebanon

Virginia Gil-Rivas University of North Carolina, Charlotte, Charlotte, NC, USA

Amanda K. Gilmore University of Washington, Seattle, WA, USA

Meyer D. Glantz National Institute on Drug Abuse, Bethesda, MD, USA; National Institutes of Health, Bethesda, MD, USA

Mark S. Gold College of Medicine, University of Florida, Gainesville, FL, USA

Vivian M. Gonzalez University of Alaska Anchorage, Anchorage, AK, USA

Mark D. Griffiths Nottingham Trent University, Nottingham, UK

Christian Grov Brooklyn College of the City University of New York, Brooklyn, NY, USA

Kristin W. Grover The University of Vermont, Burlington, VT, USA

Catherine A. Haighton Newcastle University, Newcastle upon Tyne, UK

Peter Hajek Tobacco Dependence Research Unit, Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, Whitechapel, London, UK

Wayne Hall The University of Queensland, Herston, QLD, Australia,

David J. Hanson State University of New York, Potsdam, NY, USA

Paul Harrell Columbia University, NY, USA

Deborah Hasin Columbia University, College of Physicians and Surgeons, Mailman School of Public Health, New York State Psychiatric Institute, NYSPI/Columbia University, New York, NY, USA

Penelope Anne Hasking Monash University, Clayton, VIC, Australia

Gerard Hastings Institute for Social Marketing, University of Stirling, Stirling, UK

Jennifer A. Heaton University of North Carolina Wilmington, Wilmington, NC, USA

Adrienne J. Heinz University of Illinois at Chicago, Chicago, IL, USA

Shai Hendricks Washington State University, Vancouver, WA, USA

Daniel Hermens The Brain and Mind Research Institute, Camperdown, NSW, Australia

Michie N. Hesselbrock University of Connecticut School of Medicine, Farmington, CT, USA

Victor M. Hesselbrock University of Connecticut School of Medicine, Farmington, CT, USA

Elizabeth M. Hill University of Detroit Mercy, Detroit, MI, USA

Ralph W. Hingson National Institute on Alcohol Abuse and Alcoholism, Bethesda, MD, USA

Julianna Hogan The University of Vermont, Burlington, VT, USA

Keith J. Horvath University of Minnesota, Minneapolis, MN, USA

Sharon Hsin Hsu Center for the Study of Health & Risk Behaviors, University of Washington, Seattle, WA, USA

Xiuqin Huang General Hospital of Beijing Military Region, Beijing, China

Dorian Hunter-Reel University of Washington, Seattle, WA, USA

Qiaolei Jiang The Chinese University of Hong Kong, N.T., Hong Kong SAR, China

Qiaolei Jiang Nanyang Technological University, Singapore

Thomas J. Johnson Indiana State University, Terre Haute, IN, USA

Gen Kanayama McLean Hospital, Belmont, MA, USA

Eileen F.S. Kaner Newcastle University, Newcastle upon Tyne, UK

Jon D. Kassel University of Illinois at Chicago, Chicago, IL, USA

Frances Kay-Lambkin National Drug and Alcohol Research Center, University of NSW, Sydney, NSW, Australia; Center for Brain and Mental Health Research, University of Newcastle, Callaghan, NSW, Australia

Sharlene S. Kaye Centre for Research on Ageing, Health and Well-being, The Australian National University, Canberra, ACT, Australia

Debra Kaysen University of Washington, Seattle, WA, USA

Eva Kemps Flinders University, Adelaide, SA, Australia

Jason R. Kilmer Center for the Study of Health and Risk Behaviors, University of Washington, Seattle, WA, USA

Daniel L. King The University of Adelaide, Adelaide, SA, Australia

Axel Klein University of Kent, Canterbury, Kent, UK

Jesse Kosiba The University of Vermont, Burlington, VT, USA

Jean L. Kristeller Indiana State University, Terre Haute, IN, USA

Grace P. Lee Columbia University, NY, USA

Keren Lehavot VA Puget Sound Healthcare System, Seattle, WA, USA

Michel Lejoyeux Paris 7 University, Paris, France

Carl Lejuez University of Maryland College Park, College Park, MD, USA

Kenneth E. Leonard Research Institute on Addictions, Buffalo, NY, USA

Melissa A. Lewis University of Washington, Seattle, WA, USA

Kristen P. Lindgren University of Washington, Seattle, WA, USA

Jonathan Ling Sunderland University, Sunderland, UK

Dana Litt University of Washington, Seattle, WA, USA

Andrew K. Littlefield Psychology Building, Columbia, MO, USA

M. Kathleen B. Lustyk Seattle Pacific University, Seattle, WA, USA

Laura MacPherson University of Maryland College Park, College Park, MD, USA

David A. MacQueen Moffitt Cancer Center, University of South Florida, Tampa, FL, USA

Jennifer L. Maggs The Pennsylvania State University, PA, USA

Robert Malcolm Medical University of South Carolina, Charleston, SC, USA

Christina Marel University of New South Wales, Sydney, NSW, Australia

G. Alan Marlatt University of Washington, Seattle, WA, USA

Matthew P. Martens University of Missouri, Columbia, MO, USA

Silvia S. Martins Columbia University, NY, USA

Jon May School of Psychology, University of Plymouth, Drake Circus, Plymouth, UK

Wasim Maziak School of Public Health, University of Memphis, Memphis, TN, USA; Syrian Center for Tobacco Studies, Memphis, TN, USA

Rebecca McKetin Centre for Research on Ageing, Health and Well-being, The Australian National University, Canberra, ACT, Australia

Ann McNeill Division of Epidemiology and Public Health, Clinical Sciences Building, Nottingham City Hospital, Nottingham, UK

Linda McWhorter University of North Carolina, Charlotte, Charlotte, NC, USA

Katherine L. Mills University of New South Wales, Sydney, NSW, Australia

Manish K. Mishra Dartmouth Medical School, Lebanon, NH, USA

Yamile Molina Fred Hutchinson Cancer Research Center, Seattle, WA, USA

Hamdy Fouad Moselhy UAE University, United Arab Emirates

Kim T. Mueser Dartmouth Psychiatric Research Center, Concord, NH, USA

Marcus Munafò School of Experimental Psychology, University of Bristol, Bristol, UK

James G. Murphy University of Memphis, Memphis, TN, USA

Rachael Murray Division of Epidemiology and Public Health, Clinical Sciences Building, Nottingham City Hospital, Nottingham, UK

Clayton Neighbors University of Houston, Houston, TX, USA

Nora E. Noel University of North Carolina Wilmington, Wilmington, NC, USA

Susan Nolen-Hoeksema Yale University, New Haven, CT, USA

Douglas L. Noordsy Dartmouth Medical School, Lebanon, NH, USA

Timothy J. O'Farrell VA Boston Healthcare System and Harvard Medical School, Brockton, MA, USA

Ursula Oberst FPCEE Blanquerna, Universidad Ramon Llull, Barcelona, Spain

Michael Odenwald University of Konstanz, Konstanz, Germany

Catarina R. Oliveira CNC, University of Coimbra, Coimbra, Portugal

Jason A. Oliver Moffitt Cancer Center, University of South Florida, Tampa, FL, USA

Tian P.S. Oei University of Queensland, St. Lucia, QLD, Australia

Jeffrey T. Parsons Hunter College of the City University of New York, New York, NY, USA

Megan E. Patrick University of Michigan, Ann Arbor, MI, USA

Martin P. Paulus University of California, San Diego, CA, USA

Harrison G. Pope, Jr. McLean Hospital, Belmont, MA, USA

Patricia A. Powell National Institute on Alcohol Abuse and Alcoholism, Bethesda, MD, USA

Dawn Proctor Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, QLD, Australia

Brian M. Quigley Research Institute on Addictions, Buffalo, NY, USA

Danielle E. Ramo University of California, San Francisco, CA, USA

Ana Cristina Rego CNC, University of Coimbra, Coimbra, Portugal

Jürgen Rehm Centre for Addiction and Mental Health (CAMH), Toronto, ON, Canada; University of Toronto, ON, Canada; Institut für Klinische Psychologie und Psychotherapie, TU Dresden, Germany; Dalla Lana School of Public Health, Toronto, ON, Canada

Elizabeth K. Reynolds Johns Hopkins School of Medicine, Baltimore, MD, USA; University of Maryland College Park, College Park, MD, USA

Jennifer A. Robinson PsycHealth Centre, Winnipeg, MB, Canada

Lindsey Rodriguez University of Houston, Houston, TX, USA

John M. Roll Washington State University, Spokane, WA, USA

Abigail K. Rose University of Liverpool, Liverpool, UK

Harold Rosenberg Bowling Green State University, Bowling Green, OH, USA

Desiree M. Rudisill Washington State University, Vancouver, WA, USA

Andriy V. Samokhvalov Centre for Addiction and Mental Health (CAMH), Toronto, ON, Canada; University of Toronto, Canada; University of Toronto, ON, Canada

Zila M. Sanchez Universidade Federal de São Paulo, CEBRID (Brazilian Center of Information of Psychotropic Drugs), São Paulo, SP, Brazil

Natalia Santoya Washington State University, Vancouver, WA, USA

Rebecca L. Schacht Johns Hopkins University School of Medicine, Baltimore, MD, USA

Trevor J. Schraufnagel University of Washington, Seattle, WA, USA

John E. Schulenberg The University of Michigan, MI, USA

Jeremiah A. Schumm Cincinnati VA Medical Center and University of Cincinnati, Cincinnati, OH, USA

Lion Shahab Health Behaviour Research Centre, University College London, London, UK

Saul Shiffman University of Pittsburgh, Pittsburgh, PA, USA

Richard L. Shriner College of Medicine, University of Florida, Gainesville, FL, USA

Rajita Sinha Yale University School of Medicine, New Haven, CT, USA

Monica C. Skewes University of Alaska Fairbanks, Fairbanks, AK, USA

Andrew Paul Smith Centre for Occupational and Health Psychology, School of Psychology, Cardiff University, Cardiff, UK

Karen E. Smith Sunderland University, Sunderland, UK

Marcantonio Spada London South Bank University, London, UK

Alan W. Stacy Claremont Graduate University, Claremont, CA, USA

Gillian Steckler Washington State University, Vancouver, WA, USA

Jennifer L. Stewart University of California, San Diego, CA, USA

Carla L. Storr University of Maryland School of Nursing, Baltimore, MD, USA

Esben Strodl Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, QLD, Australia

Tian Po Sumantri Oei The University of Queensland, St Lucia, QLD, Australia

Ran Tao General Hospital of Beijing Military Region, Beijing, China

Maree Teesson National Drug and Alcohol Research Centre, University of New South Wales, Sydney, NSW, Australia

Jo Thakker University of Waikato, Hamilton, New Zealand

Judy Tidwell University of Houston, Houston, TX, USA

Marika Tiggemann Flinders University, Adelaide, SA, Australia

J. Tirapu-Ustarroz Clínica Ubarmin, Egüés-Navarra, Spain

Tony Toneatto University of Toronto, Toronto, ON, Canada

Lynn M. Veatch Medical University of South Carolina, Charleston, SC, USA

Jennifer C. Veilleux University of Illinois at Chicago, Chicago, IL, USA

A. Verdejo-García Institute of Neuroscience F. Olóriz, Universidad de Granada, Granada, Spain

Nasir Warfa Wolfson Institute of Preventive Medicine, London, UK

Aviv Weinstein Hadassah Medical Organization, Jerusalem, Israel

Robert West Health Behaviour Research Centre, University College London, London, UK

Joseph Westermeyer University of Minnesota, Minneapolis, MN, USA; Minneapolis VA Medical Center, Minneapolis, MN, USA

Aaron M White National Institute on Alcohol Abuse and Alcoholism, Bethesda, MD, USA

Reinout W. Wiers University of Amsterdam, Amsterdam, The Netherlands

Graeme B. Wilson Newcastle University, Newcastle upon Tyne, UK

Rachel P. Winograd Psychology Building, Columbia, MO, USA

Adam R. Winstock South London and Maudsley NHS Trust, Kings College London, London, UK

Katie Witkiewitz University of New Mexico, Albuquerque, NM, USA

George E. Woody University of Pennsylvania and Treatment Research Institute, Philadelphia, PA, USA

Laura Wray-Lake Claremont Graduate University, CA, USA

Li-Tzy Wu Duke University Medical Center, Durham, NC, USA

Ross McD. Young Institute of Health and Biomedical Innovation, Queensland University of Technology, QLD, Australia

Michael J. Zvolensky The University of Vermont, Burlington, VT, USA

Ivori Zvorsky University of Richmond, Richmond, VA, USA

CHAPTER

1

Historical Understandings of Addiction

Joseph Westermeyer[*,$]

[*]University of Minnesota, Minneapolis, MN, USA

[$]Minneapolis VA Medical Center, Minneapolis, MN, USA

OUTLINE

BACKGROUND

Rationale for Addiction Terms and Concepts

The term addiction, from the Latin addictionem or a devoting, first appeared in the English language around the sixteenth century. During recent historical times, words that indicate alcoholism or opiate addiction have appeared in many languages. Even in societies without written language, such terms exist. For example, in the Tai-Lao languages of Southeast Asia, these terms involve one word meaning attached to or stuck to be followed by the term for alcohol or opium (e.g. teat-dyafeen, literally stuck, attached, or connected to opium for opium addiction). At times, these terms have a derogatory connotation. For example, in the Tai-Lao languages, a deteriorated person dependent on alcohol may be called khun-kee-lao or literally person-excrement-alcohol. A person who spent most of the day intoxicated on opium might be referred to as khun-kee-dyafeen or person-excrement-opium.

The populace at large uses these terms commonly in societies with widespread alcohol or drug addiction. They are not technical terms employed only by healers, jurists, or literati. Their popular usage indicates that addiction occurs sufficiently often to warrant a special term. When enough people manifest addiction, a term has evolved to ease communication about them, their behavior, or their own personal experience (e.g. "When I became ill after I stopped using, I realized I was addicted").

These special terms foster abstract thought about addiction. Not surprisingly, models have appeared

http://dx.doi.org/10.1016/B978-0-12-398336-7.00001-2

over the centuries to explain or elaborate on the experience or observation of substance use. Related concepts have also evolved, such as ideas regarding causation, related manifestations, or solutions to addiction-related problems. As with all terms in common lay usage, they involve not only denotative or cognitive aspects but also connotative or value-laden aspects. For example, a homeless alcohol-dependent patient wanted me to know that he considered himself a common drunk, but under no circumstances was he an alcoholic.

Models for Episodic Intoxication in Traditional Societies

All cultures have forms of time out in which the daily customs or rules governing roles, responsibilities, and behavior are temporarily in abeyance. These are often time-limited periods of celebration, socialization, relaxation, and/or feasting. For example, Roman festivities involved wine drinking with feasting on special foods. The so-called "drunken comportment" may itself signal respite from one's usual roles and responsibilities.

In some cultures, substance use may involve a highly valued altered state of consciousness. The latter states may provide a means of communing with spirit world, so that individuals may be guided, inspired, or instructed on how to proceed with their lives. For example, in the Americas, aboriginal cultures often employed psychoactive substances to commune with the spirit world. Other methods, including fasting, sleeplessness, and isolation from other people, were used in the quest of supernatural guidance. In addition to ceremonial use at important annual ceremonies (e.g. Passover, New Year), psychoactive substance use has attended numerous rituals in the life of individuals. Examples include important social contracts (e.g. the du-tsen in Germanic cultures, sheltering a stranger in the Middle East, commercial or political relationships in Southeast Asia, life milestones such as graduation, marriage, childbirth, or death).

Medicinal use of psychoactive substances has long been used to relieve illness, or at least the symptoms of illness. For example, opium can relieve acute pain, diarrhea, cough, fear, dysphoria, and other forms of misery. In areas where alcohol was distilled, people added various herbal compounds to alcohol (including opium) for their presumed medicinal properties. Sometimes healers added objects to alcohol (e.g. snakes, insects, minerals) to bring the presumed special attributes of these objects to the suffering individual.

Preparation of beer, wine, and distilled beverages permitted the storage of carbohydrates in the form of alcohol. Used in this way as a foodstuff, beverage alcohol could augment the diet during times of the year that food was not being produced. These drinks introduced tastes that could enhance the taste of food. Beverage alcohol could substitute water in settings where potable water was not readily available (although excessive doses purge body fluids through its diuretic effect). Small doses of some psychoactive substances, such as cannabis, have long been used as condiments in soups and baked goods.

Use of alcohol and drugs served several religious and spiritual functions. During religious ceremonies and rituals, priests and sometimes devotees at large consumed sacramental alcohol and drugs. The latter included consumption of hallucinogenic mushrooms and other plant products (such as peyote) in the Americas. The Native American Church in North America has continued this practice in modern times as a means of communicating with the spiritual world, addressing conflicts and personal distress, and finding one's way into a moral, productive future.

Societal control predominated over individual choice in the traditional use of these psychoactive substances in earlier times. For example, all individuals might have their beverage containers filled at the same time so that no one drank more (or less) than a socially prescribed amount. Occasions of use were socially determined, limiting the frequency of use. This traditional form of control began eroding when masses of people received beverage alcohol as payment for work in industrialized England. People began to drink daily, sometimes even before starting work in the morning. This daylong drinking favored individual choice over drinking, as well as physical dependence on alcohol.

Intoxication Gone Awry: Appearance of Widespread Addiction

Addiction required daily consumption. Such use led to increasing dosage to obtain the desired effects for certain psychoactive substances, including alcohol, opium, cannabis, sedatives, and tobacco. Increased frequency and dosage almost inevitably produced consequences: economic, psycho-physiological adaptation, and biomedical complications. Depending on the substance, these consequences occurred within a few years to a decade or longer. For example, drugs such as heroin and cocaine can produce consequences sooner. Substances such as alcohol and opium require several years to a decade or longer in the average case.

A common manifestation with most psychoactive substances is the need for increasing doses to produce the same, desirable effects. Thus, an alcohol drinker who experienced relaxation and other positive effects from 1 or 2 oz. of alcohol will need twice, then four, and later ten times that amount to produce the same desired effects. The same occurs with opium and heroin, with even higher multiples in some cases. Similar escalations occur with

tobacco, cocaine, and other stimulants, and certain other drugs (e.g. sedatives, cannabis, phencyclidine).

Economic and technical advances have permitted the expenditure of time, resources, and wealth to permit widespread addiction. In past centuries, much effort was required to produce sufficient carbohydrate, opium poppy, or tobacco to permit many people to engage in daily use of large doses. In the Caribbean, slaves on plantations grew sugarcane used to distill rum and gin. Trading ships from England distributed goods along the eastern coast of North America. Before heading back home to England, they picked up ballast in the form of beverage alcohol. Likewise, the English raised poppy in India to use in trade with Southeast Asia, China, and Japan. Thus, the worldwide phenomenon of widespread addiction after 1600 AD depended on such innovations as the following:

- Production of excess carbohydrate in large amounts, as a result of slave labor, or plantation-based mass production, and/or technical innovations (such as the iron plow, use of fertilization, irrigation) to produce alcohol.
- New methods of growing labor-intensive drugs, such as opium, tobacco, or betel-areca nut, again using plantations and new agro-technology.
- Distillation of beverage alcohol and drug preparations to extract active agents (e.g. morphine or heroin from opium), which reduced the cost of transportation over great distances, making lower cost alcohol or drugs available in areas that had no experience, and hence no resistance to their use.
- The short half-life of some purified substances (e.g. heroin) as compared to the parent compound (e.g. opium) hastened the development of addiction as well as the consequences wrought by addiction.

MODELS FOR UNDERSTANDING ADDICTION EXISTING PRIOR TO 1600 AD

People cannot avoid positing causes for observed behaviors, especially when these behaviors are problematic. Inevitably, these causal explanations, also called models or paradigms, tend to lead to interventions.

Moral Model

In many societies after 1600 AD, citizens no longer hued to one tradition governing use of psychoactive substances. Organized religion often fed this diversity, as adherents of one sect forbade use while those in other sects approved moderate use. Eventually, decisions whether to use, when to use, and how much to use fell to individual choice in many times and places. In turn, this led to presumed morality in making such choices. Abstention or moderate use was deemed as morally good, whereas excessive, problematic use was judged as morally bad. Those consistently making bad decisions came to be viewed as morally weak, corrupt, or full-of-sin (sinful).

This model prevails today in many instances. For example, if a society values individual responsibility and accountability, people will view many, if not most of their important decisions as having moral implications. Purchase and consumption of alcohol in significant amounts can deplete families of necessary resources, result in poor health that drains family well-being, and led to irresponsible behavior that harms family members or other members of society.

As addicted persons fail in their attempts to change or curb their own behavior, they often apply this model to themselves. That is, they perceive themselves as morally bad or having weak character. Families and friends may also apply a moral model in judging the consumptive behavior of an addicted person. The moral model may lead to changes in behavior that can be successful in early stages of heavy use or addiction. Later, these models seldom work for any length of time. Most clinicians learn that this model seldom helps in the case of severe addiction.

Organized religions have addressed excessive substance use and addiction as a moral dilemma for their adherents. Many religions have prescribed total abstention as a solution. For example, Hinduism, Buddhism, and Islam have opposed the use of alcohol and recreational drug use for centuries. Although some use may be permitted in some Hindu, Buddhist, and Islamic societies, the clergy and the religious laity have been urged toward total abstinence. In Judaism and Roman Catholicism, the use of alcohol in religious rituals has imparted a message that this sacramental substance should be respected and not abused. Following widespread addiction in the seventeenth and eighteenth centuries, many abstinent-oriented Christian and other sects arose, forbidding use among their members.

Criminal Model

The criminal model bears a strong relationship to the moral model, from which it stems. As societies experienced increasing social problems associated with addiction, some societies defined addictive behavior or excessive alcohol–drug consumption as a crime against society, and not simply a moral failing. Depending on the society and the point in time, any public intoxication might be punished. Or only dangerous behavior associated with intoxication (such as fighting or beating family members) might be punished.

In colonial times in the United States, heavy or addicted drinkers might be placed in wooden stocks, located in a public place. Not only would this punishment restrict the addict's freedom, it would also impose discomforting immobility in the stocks. Since stocks were typically located in a public place, the punishment involved public shaming. Whether it proved effective in ameliorating or reducing addiction is unknown, but it did create a public attitude against addiction.

This model still persists in many venues today. For example, some communities publish the names of those arrested for drunken driving or public intoxication. Pregnant mothers whose use of alcohol or drugs may threaten the development of their baby in utero may have their freedoms circumscribed until childbirth. The war on drugs policy in many modern nations, in which drugs are viewed as foisted on innocent people by criminal producers and traffickers, owes its persistence to this model.

Preternatural Model

In many cultures, alcohol and drugs were imbued with preternatural or even supernatural powers. The preternatural model also owes its origins partially to the moral model, as people believed that no rational person would choose addiction over abstinence of moderate use. That is, the psychoactive substance was viewed as having the ability to take over the person's mind, will, or soul through means that transcended the natural world. This model has sometimes been called the devil in the drug model. In this view, substances such as alcohol, opium, or cannabis are so inherently evil and even demonic that any use is evil. Such use may then be labeled as sinful, so that the person's inherent goodness may be doubted or denied. Although this model is still popular in many settings, increasing knowledge of the neurotransmitter, neuroanatomical, and genetic nature of addictogenic psychoactive substances has challenged these early supernatural explanations, while also providing explanations regarding why alcohol and drugs can weaken the will and moral resolve of addicted persons.

MODELS APPEARING BETWEEN 1600 AND 1900 AD

Epidemic Model

Two epidemics of addictive disorder appeared in the 1600s and spread notably by the early 1700s. They occurred around the same time, although they involved different substances in different parts of the world. Moreover, they both involved a complex of new behaviors not modified by tradition, plus international trade.

The English gin epidemic (which also involved rum) involved the import of cheap distilled beverage alcohol from the Caribbean area. In the early stages of the Industrial Revolution, English ships brought merchandise to Canada, the American colonies, and then the Caribbean. Needing ballast for the return trip to England, they first used stones, then sugarcane and other agricultural products, and finally distilled alcohol. On the docks of Liverpool, a calorie of beverage alcohol could be purchased for less than calorie of bread. Factory owners often paid workers in alcohol. Unfettered by the traditions governing use of mead, workers made their own decisions regarding when and how much to drink. The result was widespread excessive drinking.

Other places in Europe as well as North America similarly became enamored with heavy drinking. Led by England, several responses gradually led to a reduction in alcohol addiction. One of these interventions involved taxing alcohol beverage. Abstinence-oriented Christianity began around this time. In England, popular eight-page booklets described the depredations of alcohol, together with wood-block pictures of the consequences, including fetal alcohol syndrome, child neglect and abuse, fights, accidents, theft, and poor workmanship. In the United States, small residential asylums were developed to treat chronic alcohol abuse.

Within years after contact was established between the Americas and Eurasia, tobacco smoking appeared in the Orient. Tobacco smoking houses appeared, frequently by youth, political dissidents, and other independent-minded people. Several nations of Asia closed these public smoking houses. The result was the wedding of opium, an old substance once consumed only by eating, and smoking, a new method of administration. Opium smoking spread like wildfire in Southeast Asia and the East Orient, including China, Korea, and Japan. As in Europe, many strategies against opium addiction began. One of these was anti-opium societies, whose members endeavored to bring addicted people to residential facilities that provided withdrawal treatment, nourishment, and respite. Several countries also passed laws against the production and/or import of opium. In China, a cabinet-level minister was appointed to limit production, commerce, and use of opium. The United States mimicked many of these measures over time.

Illness Model

As noted above, this model evolved in Asia, Europe, and North America to counter widespread opium and alcohol addiction. During that period, it typically

involved admission to a hospital or residential facility, often in a rural area. So-called moral treatment, developed for psychiatric disorders, included shelter, nourishment, daily activities, supervision, and respectful social interactions.

This model is often inappropriately described as the medical model of addiction. Although the term medical model may apply to the downstream attempts to reverse addiction in medical settings or by physicians, it is the illness model that gave rise to treatment in the first place. This model evolved over two centuries ago when addicted persons were unable to cease addictive use of psychoactive substances on their own.

One of the first instances of this model occurred in the United States around 1800. Small therapeutic households (or asylum) developed in rural areas to provide a period of supervised abstinence for alcohol addicts. Although some physicians were active in developing this resource, nonphysicians established and ran these recovery-oriented households.

During the Opium Epidemic in China and other parts of Asia, small clinics and hospitals were established to support addicted persons through withdrawal and a return to health. Locally established anti-opium societies provided material support and referred addicts to these facilities.

A social concern related to this model has focused on the fear that addicts might present the illness model as a rationale for continued addictive use of a substance. Although this does occur rarely in addicted persons who may want to continue their addictive use, most addicted persons develop a moral imperative to return to responsible living, once they have some months of healthful sobriety. Another fear relates to the worry that unethical clinicians might provide ineffective treatment as a means of accumulating wealth. Again, this can occur; but ethical clinicians do not experience enhanced self-esteem in their healing roles by sponsoring ineffective therapies.

The illness model involves processes that proceed along psychological, psychosocial, and cultural lines. First, the addicted person must view himself or herself as a blighted or diseased person in need of outside help. This step, involving illness behavior, occurs after a period of misery and dysfunction. Second, those around the person must be willing to deed the person a period of relief from ordinary social expectations and responsibilities, in order to permit treatment and recovery. This involved social assignment of a temporary sick role. Third, a culturally approved or licensed health care worker must ordain that disease exists and treatment is warranted. The social assignment of a sick role usually continues as long as the suffering person is willing to undergo the physical, psychological, or social burdens associated with treatment.

MODELS APPEARING SINCE 1900 AD

Personality Disorder Model

During the early 1900s, diagnostic classifications included alcoholism and drug addiction as a personality or character disorder. These disorders were viewed as a form of antisocial personality, since the individual broke social mores, acted primarily in their own apparent self-interest, and often transgressed the rights of others while intoxicated or drug seeking. According to this view, addiction evolved in irresponsible or self-centered people who ignored the effects of their choices and behaviors on others. The second edition of the Diagnostic and Statistical Manual of the American Psychiatric Association typified this perspective.

During this period, personality and character disorders were seen as untreatable. Thus, this model justified the noninvolvement of many clinicians in the care of these patients. Despite this professional viewpoint, many states provided asylum-type care. Private asylum or dry-out farms also operated.

Abandoned by the medical profession, alcoholics in the United States supported one another's recovery in the brotherhood of Alcoholics Anonymous (AA). One of the founding members of AA, himself a physician, used the analogy of allergy to describe the individual, and presumed idiosyncratic response of alcoholics to alcohol. This analogy bears strong resemblance to the Illness Model from the previous century.

Alcoholics Anonymous borrowed heavily from an English self-improvement group that aimed at maturity and balance, using early Christian principles. Although not sponsored by organized religion, it employed spiritual guidance and growth as a means of achieving a rewarding, responsible life way in the modern industrialized world.

Neurotransmitter Model

Alcohol, opium, and other substances of abuse mimic or affect naturally occurring neurotransmitters. Neurotransmission was key to understanding pain and analgesia, anxiety and relaxation, dysphoria and euphoria, dissociation and vigilance. Increased understanding of drug/alcohol-related effects on brain and behavior has led to other theories, as well as to new therapies of addiction. In turn, new information accruing from successful treatment of addiction has enriched our theoretical perspective on addiction.

Various drugs of abuse relate to specific neurotransmitter systems. For example, opioid drugs mimic the effects of endogenous opioid-like substances (i.e. endorphins). Cocaine, amphetamines, and other stimulants affect the adrenergic and dopaminergic systems.

Nicotine affects muscarinic receptors. Cannabis and its active compound tetrahydrocannabinol stimulate an endogenous cannabinoid receptor. Benzodiazepines and other sedatives trip a benzodiazepine receptor, which in turn affects the actions of gamma-aminobutyric acid. Medications anatagonistic to these substances (e.g. disulfiram for alcohol, naltrexone for opioids) have been used in treatment.

Neuroanatomic Model

Certain areas of organelles of the brain are related to alcohol or drug effects. For example, the dopaminergic locus accumbens appears to be a site involved in the reward experienced with psychoactive substance use. Alcohol, sedatives, and opioids reduce the contributions of the frontal lobes that facilitate recall, executive decision-making, interpersonal sensitivity, judgment, and morality.

Awareness of these brain functions and their dysfunction during addiction contributes to an understanding of why addictive behaviors may persist despite their damaging effects on the individual. Damage of particular areas of the brain can aid clinicians in recognizing certain stereotypic conditions associated with addiction. For example, damage to the mammillary body area of the brain can produce a chronic inability to store recent memory, a lesion that produces an inability to live independently. Damage to the frontal lobes can foster disinhibited, intrusive speech and behavior.

Learning Model

This model views addiction as a learned behavior that follows upon certain learned stimuli. For example, a person might experience withdrawal symptoms in a setting where alcohol or drug use had previously occurred. In this example, the effect of the alcohol or drug use is the unconditioned response and the alcohol or drug itself is the unconditioned stimulus. The environmental cues associated with the presence of alcohol or drugs are the conditioned stimuli that can produce drug seeking and drug using.

This model can be used therapeutically by avoiding cues that stimulate urges to use and by extinguishing craving responses to these cues. Recalling and imaging the aversive consequences brought on by use may also be employed. Medications that eliminate the desired drug or alcohol effects can be useful in extinguishing addiction-related learned behavior. On the level of the community, learning can be used to reward behaviors inconsistent with alcohol or drug abuse (e.g. family interactions, employment, abstinent-oriented recreation).

Dyadic Enabling/Rescuing Model

Described first by Eric Berne, this model involves the interaction between at least two people. One of these could be an addicted person, who relates to the second person in a child-like or dependent role. The second individual relates to the addicted person as a parent, helper, or authority figure. This role relationship is apt to accentuate the dependent behavior of the addicted person, while frustrating the second person who expects the proffered help to ameliorate the addiction (which it seldom does, if given without contingencies).

The helper may enable the addictive behavior by providing shelter, food, resources, or even drugs or alcohol to the addicted person. Enabling fosters continued using alcohol or drugs without experiencing the social, economic, or other consequences of use (e.g. poverty, no food or shelter). The ultimate effect is a worsening of the addicted person's condition.

Rescuing involves the parental, helping person saving the child-like, dependent person from social or other consequences of addiction. For example, a policeman may cite a drunk driver for another, lesser offense than the actual offense. Or a judge might not exact a legal punishment because it is a first offense. Rescuing, like enabling, is often undertaken in hopes that the beneficent action will goad the addicted person toward recovery. Unfortunately, rescuing generally exacerbates rather than alleviates addiction.

Route-of-Administration Model

Many substances can be consumed by more than one route of self-administration. Routes that have utilized over time include the following:

Route	Absorption	Rapidity of effect
Eat/drink	absorption	slow (10 to 20 minutes)
Subcutaneous	skin tissue (injection)	slow
Intramuscular	muscle tissue (injection)	slow to medium
Chewing	oral mucosa	medium (several minutes)
Rectal	rectal mucosa	medium
Vaginal	vaginal mucosa	medium
Snuffing	nasal mucosa	rapid (< one minute)
Smoking	pulmonary alveoli	very rapid (< 10 seconds)
Intravenous	blood stream (injection)	very rapid

Biomedical consequences may be linked to the mode of administration, to some extent independent of the substance used. For example, alcohol drinking can cause ulcers or cancers. Tobacco, cannabis, or opium smoking can produce chronic lung infections and airway/pulmonary cancers. Chewing astringent substances, such as tobacco or betel nut, can cause gum inflammation, dental lesions, and oral cancers. Snuffing with cocaine or heroin can produce ulceration and even perforation of the nasal septum. Subcutaneous, intramuscular, or intravenous injection can produce bacterial, fungal, or viral infections. More rapid routes of administration tend to be more addicting, especially if the onset-of-action of the drug is rapid.

Comorbidity Model

This concept suggests that all or much substance use occurs in association with other comorbid conditions. Some of these conditions may occur before the onset of substance use disorder and conduce toward subsequent substance use disorders. This could occur if chronic pain led to excessive opioid drug use, or chronic insomnia or anxiety led to excessive use of sedatives. On the contrary, substance use disorder could lead to comorbid conditions. For example, losses associated with substance use disorder might precipitate depression. Or various routes of administration (see above) might favor the development of biomedical complications.

Virtually any psychiatric disorders can accompany substance use disorders. However, mood, anxiety, and other externalizing disorders are most frequent psychosocial disorders. Psychosis and brain injury occur more often than chance. Numerous medical problems can accompany substance use disorder. These also include various infectious, gastrointestinal, cardiovascular, genitourinary, and traumatic maladies. Alcohol involves several unique metabolic and nutritional problems.

Salutogenic or Self-treatment Model

Drug or alcohol use may be used for various salutogenic purposes. For example, a person may use a substance to feel relaxed in a social situation, to augment sexual experience, or to remain alert when fatigued. Khantzian extended this model to include self-treatment for discomforting or even disabling symptoms. For example, an anxious person might use sedative or opioid drugs or alcohol to relieve chronic anxiety, panic, or phobias. A depressed person may use cocaine or amphetamine to relieve fatigue or lack of concentration.

With time, the drug may have an effect opposite to that intended. For example, opium might enhance sexual function early in its use but ultimately cause loss of libido and impotence. Or alcohol might facilitate social interaction at the beginning but lead to social alienation and isolation with excessive abuse. Moreover, more than one type of drug could be used for a specific disorder. For example, sedatives might be used to relieve insomnia, stimulants to alleviate fatigue, and opioids to relieve a lowered pain threshold.

Genetic Model

The genetic model posits that genes either increase or decrease one's vulnerability to alcohol or drug addiction. Studies of twins and adoptive studies both confirm that genetic inheritance contributes to the development of addiction. Having two parents with addiction further contributes to the severity of addiction. Nonetheless, up to 40% of patients presenting for treatment of addiction do not have a parent with an addictive disorder (but may have a grandparent or other relative with substance abuse). Genetic inheritance may not be a sine qua non for addictive disorders, or genetic vulnerability may not be manifest without certain environmental factors. Recent data suggests that disruptive environment during childhood or adolescence interacts with genetic vulnerability to increase the risk of addiction.

Externalizing Disorder Model

Child–adolescent psychologists and psychiatrists first described externalizing disorders as those involving behavior, including conduct disorders, substance use disorders, pathological gambling, and other behavioral problems. Internalizing disorders consisted of mood, anxiety, and somatic disorders. More recently these concepts have been applied to adults.

This model suggests that a person with any one externalizing disorder is prone to develop another externalizing disorder at some point. For example, an individual recovered from substance use disorder may be at risk for pathological gambling. The application of these concepts in adults remains in its early phases at this time, but early work suggests that this model may be important in ameliorating the longitudinal development of comorbid disorders, among those who already have had a psychiatric disorder of some sort.

Recovery Model

Treatment alone by clinicians cannot alleviate addiction. The active, committed participation of the addicted person is critical to a successful outcome. This patient-focused aspect of the treatment process is often referred to as recovery.

Recovery occurs over time in stages, rather than in a single, sudden step. The first stage of recovery involves breaking from addictive use. The goal of this phase is safety from continued addictive use, with treatment of associated biomedical and psychosocial disorders. Under optimal circumstances, this phase may last several weeks.

The second phase involves stability, including shelter, daily structure, and social support in a setting that reduces continued exposure to alcohol or other drugs of potential abuse. Recurrence of addiction is most apt to occur during this period of increasing comfort, independence, and growing self-esteem. This phase may be as short as several months, but can last longer.

The third stage of recovery involves a return to a comfortable, responsible, and productive life way. The risk to recurrence is present but decreases once the person has achieved 2 years of sobriety. This phase, if successful, requires several years or longer to become firmly established.

This eventual self-actualization associated with this model may lead relatives, clinicians, and various social gatekeepers to insist that the addicted person voluntarily seek recovery. However, the early steps toward recovery are not always so simple. Coercive treatment can help in motivating the initially unmotivated patient. Early on, coercive abstinence or treatment may aid the patient in joining the recovery pathway. In recent years, so-called drug courts have facilitated many addicted persons in turning from a life of property crime and family irresponsibility to a sober, productive life style.

Readiness-to-Change Model

Prochaska and DiClemente have described the addicted person's attitude toward recovery during the process from pre-recovery to full recovery. The first attitudinal stage, the so-called pre-contemplative stage, does not involve any intent to quit substances.

As problems mount, the addicted person may come to realize that continued addiction is causing these problems. The clinician may inform the person that addiction is causing the person's growing problems. These insights may lead to the second contemplation stage, in which the addicted person begins thinking about decreasing dose or frequency of use, or quitting use altogether. Growing ambivalence toward drug or alcohol use predominates at this stage.

The third step involves a decision to stop use, as the ambivalence toward psychoactive substance abuse crystallizes in the direction of reducing the amount and/or frequency of use, or perhaps the type of substance being used. Should this cut back approach fail (as it typically does), the addicted person may decide to seek treatment. The treatment-seeking stage requires collaborating with treatment in achieving safe, then stable, and eventual comfortable recovery.

Finally, continued maintenance of recovery builds a flexible resistance against recurrence. This phase may involve devotion to others, the environment, or some cause or purpose that serves a greater good. It may involve avoiding situations that precipitate craving. Avoiding excessive fatigue, hunger, or stress can also gird the recovering addict against relapse.

The Harm Reduction Model

The harm reduction model has evolved as a public health approach to widespread alcohol and drug addiction. The strategy lies in reducing the consequences of drinking or drugging, while assuming that alcohol and drug addiction will continue. For example, a wet house may permit inhabitants to drink heavily in the privacy of their rooms, so long as they do not aggress others or wander out into public arenas. Similarly, drug sales and injection may be provided in a safe setting, where unsafe injection, robbery, or assault can be prevented. Provision of clean needles to prevent HIV infection is another example.

In some cases, repeated attempts at treatment may end in failure and demoralization. Some addicted persons may give up the difficult struggle required for a successful recovery. For these patients, various harm reduction approaches have been employed. For example, clinicians who provide information to the patient about the harm issuing from their continued use may help the person to reduce the use. Or residence in a quarter-way house in which quiet, private drinking or drug use is permitted. Assessment of an early clinic that tried this approach revealed a high mortality rate over a 2-year period, but with eventual abstinence in a notable number of patients.

At times some social programs appear to follow a harm enabling approach. Providing shelter, income, or other resources that are not tied to a reduction in harm to self and others can increase rather than reduce harm. For example, a pension, trust, or savings can facilitate the continued deterioration of an addicted person. A legally assigned payee can route these funds to pay for shelter, food, clothes, and recovery-oriented expenditures.

CONCLUSION

No one model provides a complete understanding of all aspects of alcohol and drug addiction. A model that may help us in guiding our own use (or nonuse) of substances, such as the moral model, may not help us

in coping with an addicted relative, friend, client, or patient. An illness model may be useful for the addicted person willing to enter treatment, but not help a person who wants to continue addictive use of substances.

Although this list of models for understanding addiction is not exhaustive, nonetheless the list of models above shows the accelerating development of models for helping us to understand addiction. For example, in all the centuries before 1600 AD, only a limited few models were available to help people understand addiction. Major examples included the moral, criminal, and preternatural models.

Then following the first appearance of the widespread alcohol and opium addiction, health-related models appeared. These included the epidemic model and then, around the same time, the illness model of addiction. These new models appeared over a relative brief period of only a few centuries, greatly increasing the rate of new model development for understanding addiction. These new models did not eliminate the earlier moral, criminal, and preternatural models, but they provided useful new models for understanding and acting upon addiction. Diagnostic typologies and novel modes of treatment evolved from these models.

Over the last century, the number of new models has increased at a rate vastly greater than the rate preceding it. These new models have evolved from increased knowledge regarding learning, pharmacology, neurophysiology, and neuroanatomy. Cogent clinical observations have led to our appreciating the mode of administration, the interactions and sometimes disconnections between treatment and recovery, and the coercive role of society in requiring sobriety under certain circumstances. Randomized controlled clinical trials have increased our knowledge of what interventions work, when, and with whom. This knowledge has increased our understanding of the addictions.

Other and newer models will hopefully evolve in the coming decades to aid us in further understanding addiction. The centuries-old quandaries posed by addiction have continued to challenge addicted individuals and their families, along with the best minds in our communities. The history of these developing models has shown that enhanced understanding of the drug–person–society triad ultimately leads to effective prevention, early intervention, treatment, and rehabilitation.

SEE ALSO

The Biopsychosocial Model of Addiction, Disease Model, Self-Medication, The Terminology of Addictive Behavior, Models of Relationships between Substance Use and Mental Disorders, Minority Groups and Addictions, Spirituality and Addiction

Glossary

AA Alcoholics Anonymous.
Addictogenic substance a psychoactive substance whose use can produce dependence.
Ballast heavy weight stored low in hold of sailing vessels to prevent the vessel from tipping over in a strong wind.
Ceremonial or ritual drinking the use of alcohol takes on shared social or cultural symbolism (e.g. interpersonal, celebratory, or spiritual meanings) beyond the mere consumption of beverage alcohol.
Contingencies rewards or punishments that ensue from an addicted person's behavior.
Du-sten a German ritual in which two people agree to address one another by the informal "du" rather than the formal "Sie" when conversing together.
Endogenous a substance that is produced by the body.
Epidemic high prevalence of disease or disorder, spreading from one person or group to others, so that an entire society may suffer its ill effects.
Fetal alcohol syndrome maternal drinking causes damage to the developing fetus, so that the newborn has mental retardation and characteristic facial configuration.
GABA gamma-aminobutyric acid, a substance that affects neurotransmission.
Half-life of psychoactive substances the time required for half of an ingested substance to be deactivated (by excretion or metabolic breakdown of the substance).
Mead an alcohol-containing beverage derived from honey.
Native American Church a pan-tribal religion practiced by American Indian people in North America; it involves elements of ancient tribal religions and Christianity, with peyote consumed as a sacramental substance as a means of experiencing communication with the spiritual realm.
Neurotransmitter a biologically active chemical that affects transmission from one nerve cell to another; addictive substances mimic or affect neurotransmission.
Passover an annual religious celebration and feast in the Jewish religion.
Peyote a hallucinogenic drug obtained from a mushroom that grows wild in the Americas.
Preternatural events occurring outside of nature, such as by magic or shamanistic powers.
Psychosis a mental disorder marked by hallucinations, delusions, and/or other profound cognitive impairments.
Salutogenic favoring good health.
Smoking the volatilization of a psychoactive substance into a gas, so that it can be inhaled into the lungs and absorbed into the blood within the pulmonary alveoli.
Stocks a punishment in which the individual's extremities, and sometimes their head, are inserted through a wooden or other structure, such that they are restrained from moving about at will.
Tai-Lao languages a related family of languages spoken by hundreds of millions of people in Laos and Thailand (where they comprise the national languages), China, Vietnam, Malaysia, and Burma.
Tradition beliefs or behaviors that persisted across generations within a group.

Further Reading

Anawalt, P.R., Berdan, F.F., 1992. The Codex Mendoza. Scientific American 266, 70–79.

Arif, A., Westermeyer, J. (Eds.), 1988. A Manual for Drug and Alcohol Abuse: Guidelines for Teaching. Plenum, New York.

DuToit, B.M., 1977. Drugs, Rituals and Altered States of Consciousness. Balkema Press, Rotterdam.

Galanter, M., Kleber, H.D. (Eds.), 2008. Textbook of Substance Abuse Treatment. American Psychiatric Press, Inc., Washington, DC.

Jilek, W.G., 1977. A quest for identity: therapeutic aspects of the Salish Indian guardian spirit ceremonial. Journal Operational Psychiatry 8, 46–51.

LaBarre, W., 1964. The Peyote Cult. The Shoe String Press, Hamden, CT.

Musto, D.F., 1973. The American Disease: Origins of Narcotic Control. Yale University Press, New Haven, CT.

Popham, R.E., Schmidt, W., De Lint, J., 2000. The effects of legal restraint on drinking. In: Kissin, B., Begleiter, H. (Eds.), The Biology of Alcoholism. Plenum Press, New York.

Rodin, A.E., 1981. Infants and gin mania in 18th century London. Journal American Medical Association 245, 1237–1239.

Westermeyer, J., 1976. Primer on Chemical Dependency: A Clinical Guide to Alcohol and Drug Problems. WIlliams-Wilkins Publishers, Baltimore.

Westermeyer, J., 1979. Medical and nonmedical treatment for narcotic addicts: a comparative study from Asia. Journal of Nervous and Mental Disorders 167, 205–211.

Westermeyer, J., 1982. Poppies, Pipes and People: Opium and Its Use in Laos. University California Press., Berkeley, CA.

Westermeyer, J., Dickerson, D., 2008. Minorities. In: Galanter, M., Kleber, H.D. (Eds.), Textbook of Substance Abuse Treatment. American Psychiatric Press, Inc, Washington, DC, pp. 639–651.

Relevant Website

Harper, D., 2010. Addiction. Available from www.etymonline.com.

CHAPTER

2

The Terminology of Addictive Behavior

Meyer D. Glantz[*,$]

[*]National Institute on Drug Abuse, Bethesda, MD, USA [$]National Institutes of Health, Bethesda, MD, USA

OUTLINE

"The beginning of wisdom is the definition of terms" ***Socrates***

"Define your terms, you will permit me again to say, or we shall never understand one another." ***Voltaire***

"Truth does not consist in minute accuracy of detail, but in conveying a right impression." ***Henry Alford, quoted out of context.***

SCIENTIFIC TERMINOLOGY

All effective communication rely on clear language and consistent meaning. However, terms that are used in a scientific context are expected to meet additional standards beyond those for words in everyday communication. The choice of which scientific terms are used and the transparency of their meanings are not trivial or semantic questions nor is the use of scientific terms just proprietary jargon intended to limit public understanding of technical communication.

Specialized technical terminology with formal, systematic definitions ideally provides a precise, standardized, documented nomenclature that goes beyond the goal of general informal understanding and colloquial communication. Effective scientific terminologies serve the purpose of ensuring that statements about scientific information, observations, and theories are explicit in meaning; concrete in applicability; and replicable in the sense that the terms will have the same interpretations and specifiable implications for all users. This consistency in the definition of terms may be achieved by a reproducible operationalization; reference to a formal classification system; comparison to a concrete metric or a standardized assessment; delineation of the necessary, sufficient, and exclusory characteristics; and/or a comprehensive listing of referents included and excluded from the term's meaning.

The highest standards for scientific terms go beyond requiring unambiguous, easily and consistently understood denotations. Scientific terms are usually part of a larger system of related terms and concepts and they often reflect a particular theory, model, or at least a particular perspective. Because of this, technical terms typically have significant connotations that include assumptions and implications. These connotations should ideally be public, identified as to their theoretical framework when possible, and preferably be empirically anchored. A discussion about terminology requires consideration of the underlying concepts of the terms in question.

Every scientific field develops its own set of terms and terminology, which often change over time as do the conceptualizations and frameworks to which they relate. At one time, stating that someone was an "alcoholic" or a "drug addict" was a moral judgment and a condemnation of the person's weak will and "dissolute character." Today, these terms more often imply

http://dx.doi.org/10.1016/B978-0-12-398336-7.00002-4

Published by ELSEVIER Inc.

psychiatric conditions and as such they reflect a change in the assumptions and larger perspective on substance use problems. The meaning of these terms may well change in the future and, if so, the evolution in the nomenclature will presumably be a productive change that reflects advances in thinking about the subject. For example, a formerly prominent term that is not often used currently is "addictive personality." The term refers to the idea that some individuals have a particular type of personality that makes them especially vulnerable to alcoholism, drug addiction, and/or other compulsive behaviors such as pathological gambling. However, decreased use of the term addictive personality was not because of its connotations or characterological emphasis but because multiple research efforts did not find any empirical support for it.

CONNOTATIONS

A common vocabulary is necessary for communication in any scientific field and the scientific method requires that terms be explicitly and consistently defined to ensure usefulness, standardized identification, and reproducible interpretability of referents and the symbols that represent them. Even when a full, general, consensual common language is not achieved, the goal is that at least for the purposes of a specific communication, all of the important terms used to communicate scientific information will have the same articulated meaning, the same denotation, and/or the same operationalized definition for all of the participants in the communication. If the meanings of the major terms in a field are not clear and consistent then effective communication about science is not possible.

This principle has long been a traditional requirement for scientific and clinical issues related to substance abuse. It is the expectation of most contemporary scientific publications and clinical documents that all major terms be unambiguously, explicitly, and operationally defined. Despite this, significant issues in terminology persist regarding even some of the most fundamental terms related to substance use problems and addiction. For the most part these issues are not simple disagreements about definitions, failures of naming conventions, or referents that have not been operationalized using concrete criteria. In other words, these are not simple semantic difficulties or linguistic confusions.

Instead, many serious problems are related to inadequacies in the clarity and exposition of the connotations and implications that give larger meaning to the terms. The problems are not primarily issues about the terms themselves but about the underlying concepts and conceptual frameworks that ultimately define them and the associations they elicit. Even when it does not seem that there is a problem with terminology, there often is and it is often belied by the appearance of consensually and even operationally defined terms.

It is controversial whether problematic behaviors such as compulsive gambling, binge eating disorders, or Internet addiction should be classified in the same category as problematic substance use. However, even limiting consideration to problematic behaviors involving the use of psychoactive substances leaves many important issues. The term "drug abuse" is a good example. Sometimes it refers to virtually any use of an illegal drug, sometimes it refers to a psychiatric drug use disorder, and sometimes it is an all-inclusive general reference to any problematic substance use including alcohol or tobacco. There is not even a consistent meaning of what is considered to be a drug in the context of the term. "Drug Abuse" is a DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, 4th edition) diagnostic disorder category that is defined by a pattern of drug use leading to particular categories of social consequences or involving use in situations where it is physically hazardous. Alternatively, the ICD-10 (International Classification of Disease, 10th edition) diagnostic taxonomy specifically avoids the term abuse and only includes a category of "Harmful Use," which is defined as persistent use resulting in specifiable physical or psychological adverse consequences. Although the concept of harmful use may seem similar to the DSM-IV concept of abuse, in fact, the ICD-10 diagnosis of harmful use explicitly excludes diagnosis based on the types of substance use-related social consequences that are inherent in DSM-IV abuse. The terms alcohol and drug abuse, dependence and addiction have no consensual or consistent meaning. In reality, outside of formal diagnostic systems such as DSM and ICD, there is no standard vocabulary for problematic substance use and even with these taxonomies there is no consensus on what the cardinal problems of substance use are. Some clinicians and researchers focus on neuroadaptively based substance dependence as the primary dysfunction underlying substance use disorders while others emphasize the severely impaired function and harmful consequences associated with persistent problematic substance use as the essential disordered dimension. Perhaps this is part of the reason that there is no consensually accepted phenotype for substance use problems.

There is consensus that the major defining characteristic of problematic substance use is persistent and/or uncontrolled substance use despite hazard or harm. But there is no consensus on the concepts, connotations, and implicit paradigms associated with many of the basic and critical terms. There is no agreement on whether the primary inimical aspects of problematic substance use are the harmful consequences, the

persistent substance use reflecting a neuropsychiatric dependence, or the antisociality and often criminality of which the substance use is one of many possible manifestations. There is no consensus on whether substance use problems are most accurately viewed from the perspective of psychiatric disorders, self-regulatory deficits, pharmacologically reinforced learned behaviors, neuroadaptive changes, brain disease, and so on.

Terms such as drug abuse, alcoholism, addiction, addictive behavior, and so on, have very different meanings and implications depending on the tacit assumptions and conceptual frameworks behind them. Even seemingly small conceptual variations can lead to significant differences. For example, a long-standing issue in the field is the extent to which the legality of the substance that is used or abused is a critical factor. Some believe that if the substance is illegal, then there is likely to be a strong component of antisocial behavior involved and the illicit "drug abuse" takes on overtones of criminality. In addition, a distinction is generally maintained at a larger social level in which alcohol and tobacco use are widely accepted. For example, it has been observed that legal substances such as alcohol and tobacco have strong commercial interests both supporting and encouraging them that they be distinguished from illegal drugs.

Taking direction from this, many early investigations of drug abuse operationally defined the concept by defining virtually any illicit drug use experience to be drug abuse. These researchers compared abstainers with almost all others without differentiating subjects' levels and types of illegal drug use. Heavy users of alcohol and/or tobacco but not illegal drugs were categorized as nonabusers because the primary criterion was use of an illegal substance. Using "any illicit drug use" as the defining criterion for drug abuse obscured critical information and, not surprisingly, the findings of these studies have proved to have little value in understanding the nature of substance use problems. Assessment of individuals' levels and consequences of substance use is now standard in research.

Nevertheless, relying on the use of an illegal substance as a defining criterion continues to obscure potentially important information and to shape social and policy decisions. While criteria of drug and alcohol use problems vary, they are generally operationally defined in epidemiological studies and are typically reported as being in separate categories. Although some findings are reported in terms of total substance use/abuse, a sum of alcohol and drug use/abuse, more fine-grained analyses are not standard. At the same time, differentiation of marijuana from cocaine or heroin use/abuse is usually limited to more specialized research. While there are numerous and often valid reasons for this, the result is a reinforcement of the view that as a group, drug users/abusers are relatively homogeneous and significantly different from alcohol users/abusers. This realistically reflects some aspects of substance use/abuse, but may be misleading in regard to others. For example, researchers and policy makers often consider changes of prevalences of alcohol and drug use/abuse as independent events and rarely discuss the possibility that a decrease in one and an increase in the other may indicate less overall change in substance use behavior than it might otherwise seem. Taking this a step further are the questions of whether the legal status of tobacco use implies less harm than the use of an illegal drug and whether some people would avoid tobacco use if nicotine dependence were referred to as a drug addiction?

Using legality as a major distinction of alcohol versus drug abuse is not indefensible. It does reflect some important aspects of substance use problems. Being aware of the tacit assumptions behind terms and making them explicit is an important and heuristically powerful step. This is because terms not only reflect the concepts behind them, they also further reify and shape thinking about those concepts.

When units of the National Institute of Mental Health were expanded to independent institute level in the United States National Institutes of Health, separate institutes were created for the study of alcohol and drug problems forming the National Institute of Alcohol Abuse and Alcoholism in 1970 and the National Institute on Drug Abuse in 1974. These both reflected and strongly contributed to the perpetuation of legality as a cardinal distinction between drug and alcohol use problems. Drug and alcohol disorder treatment and prevention programs, research initiatives, and so on, are often still separated. This compartmentalization is being questioned on a broader basis and, interestingly, the Institutes are expected to merge.

Terminologies are more than assigned labels and codified descriptors. Terms are not neutral symbols, they represent concepts and theories and they convey implicit paradigms that structure and direct conceptualizations. Many alcohol and drug abuse terminology issues are a result of fundamental disagreements about the basic nature and critical characteristics of substance abuse. No particular term is necessarily right or wrong to use nor is any particular definition correct or incorrect, true or false. Nevertheless, it matters which terms are used.

EMPHASIS

Problematic substance use is often characterized as maladaptive behavior defined by persistent or uncontrolled proscribed use of pharmacologically

psychoactive substance(s) despite hazard or harm. The particular aspect of this definition that is emphasized influences the orientation and the focus of thinking about problematic substance use and the terms that are used.

The choice of which term is chosen as the primary descriptor of alcohol and/or drug use problems often implies a particular perspective about the fundamental nature and the primary concern related to problematic substance use. While these terms have changed over time in usage and connotation, following are some of the most important of these terms and the contemporary and most prominent ideas that they emphasize:

problematic substance use: a fairly neutral term notable in that it does not emphasize compulsion, physiological processes, or consequences;
substance abuse: a term that emphasizes potential or extant harmful consequences;
harmful use: a term that emphasizes harmful consequences almost exclusively;
substance misuse: a term that contrasts appropriate and inappropriate use and may include instances of unintentional inappropriate use;
substance dependence: a term that emphasizes physiological dependence as primarily indicated by increased tolerance to the substance and withdrawal which is a state of severe deprivation when use is discontinued; there is often a secondary implication of loss of control resulting from neuroadaptation to the substance;
addiction: a term that is closely related to the concept of dependence, used interchangeably by some. Increasingly, addiction has come to emphasize neuroadaptation and chronic compulsion for substance use as a result of alcohol- and/or drug-induced changes in the brain;
addictive behavior: a term that emphasizes compulsive behavior without implications about the mechanisms that underlie the compulsions; the focus of the term is often expanded beyond the use of psychoactive substances to include a broader range of reinforcing behaviors (e.g. gambling addiction);
substance use disorder: a term that emphasizes the psychopathological character of problematic substance use viewing it as comparable to other psychiatric disorders.

The connotations of terms often go beyond the aspects of the phenomena that they emphasize. Problematic substance use terms may also reflect a theoretical framework and an implicit model of the fundamental problem and necessary target for intervention. For example, a focus on problematic substance use oriented around substance dependence not only emphasizes physiological processes but also implies that successfully addressing problematic substance use requires intervention for the physiological state of dependence. Treatments that target physiological habituation, withdrawal, and tolerance such as nicotine replacement or methadone treatment typically use terms such as nicotine dependence or opiate dependence. Alternatively, programs that stress harm reduction are more likely to rely on terms such as substance abuse or harmful use while those coming from a psychiatric perspective or considering issues of comorbid psychopathologies typically refer to substance use disorders.

The implicit paradigms of terms have additional implications. For example, describing someone as having an alcohol or drug use disease implies that their problematic substance use is an involuntary, undeserved affliction while describing someone as an alcoholic or an addict is a more stigmatizing reference suggesting, at least to some, antisocial and defective character and self-inflicted adversity.

To the extent that major terms represent concepts and theories, they also predispose thinking from the implied perspective. This is not intended to suggest an extreme straw man version of this statement and propose that particular terms inflexibly structure or even strongly direct thinking, a theory sometimes described as Linguistic Relativity or the Sapir–Whorf hypothesis. Nevertheless, thinking does tend to go reflexively to the concepts and associations suggested by a term. Further, an individual's cognitive beliefs and perspective are widely accepted as influencing experience and its interpretation and terminology plays a role in the process.

MODELS AND ANALOGIES

Recognizing the influence of the connotations of scientific terms goes beyond saying that words have a positive or negative connotative valence. For example, a discussion of unrestrained alcohol use employing terms such as ego and id will lead thinking in different directions than a similar discussion using terms such as cognitive executive function, sensation seeking, or disinhibition. Taking another example, the terms drug habit, drug dependence, compulsive drug use, and drug addiction have similar and overlapping meanings but for many people they elicit different associations and implications.

Conceptualization of a scientific phenomenon is often supported with an analogy to aid understanding. For example, basic descriptions of direct electrical current often use the model of water flowing through a pipe. Such models use an established concept to convey the nature and function of a new or hard to grasp phenomenon. Though often very useful, such analogies are approximations or metaphors. They are usually limited

and may break down or be misleading beyond a particular point. The hydraulic model of direct current only applies to basic electrical phenomena and will lead to invalid inferences if applied to other electrical phenomena such as alternating current.

Analogies play an important role in how we understand phenomena and the inferences we make. Particular terms are often associated with different basic analogies and elicit different inferences. This is another way in which terminology is influential. It is especially important in the area of substance use problems where analogies and models are as contested as basic theories. The question of which model is the most appropriate analogy for substance use problems is controversial and this not only mirrors our uncertainty and disagreement about the phenomena but has also become a forum for debate about social beliefs and policies.

Public health organizations and advocates for alcohol and drug treatment have attempted to fight social stigma, self-denigration, criminal prosecution, and moral condemnation of those with substance use problems by characterizing substance addiction as a disease. The selection of the term disease rather than a less medically associated word such as dysfunction goes beyond the goal of just reducing stigma or emphasizing the biological aspects of addiction. It is a deliberate attempt to use the influence of the term to shape perception and perspective in a positive way with the contention that alcohol and drug addiction are in essence physiologically based fundamental medical conditions that should be viewed like diabetes, cancer, or other diseases.

Although the disease model acknowledges that behavior and environment play a role in the etiology and course of substance addiction, the emphasis is on physiological and presumably, at least in part, genetic factors. The primary message of characterizing substance addiction as a disease is that, as is the case once any disease has developed, the sufferer should be treated with compassion not condemned or punished. A further implication is that whether the sick person has arthritis, lung cancer resulting from cigarette smoking, heroin addiction, or any other disease, society has some responsibility to provide care for the sick individual. As is often the case with medical models, there is some additional implication that a medical and typically pharmacological treatment is the intervention approach most likely to be effective. More recently, the phrase "addiction is a brain disease" has been used to maintain that substance addiction is neurological in nature and that addictive use of alcohol and drugs characteristically relates to neuroadaptations that involve long-lasting functional disruptions of neural circuits. The implication is that the basic mechanism of addiction impairs the individual's ability to behave in ways that would enable the individual to control the disease.

Framing alcohol and drug addiction as a disease is neither invalid nor deceptive and the focus on treatment rather than blame is commendable. The use of terminology for advocacy (or condemnation) is not unusual. Researchers and clinicians in the field of substance use problems may be forgiving of whatever bias is part of describing addiction as a disease because they are likely to agree with the intent to establish a less pejorative term and perception. Nevertheless, most clinicians or researchers would not expect to find a chapter on addiction in an encyclopedia of brain diseases. Further, some express reasonable concerns that using a disease metaphor de-emphasizes behavioral, psychological, social, and environmental aspects of substance use problems.

DIAGNOSTIC TERMS

The context in which a term is used and the purpose of the terminology can have an important impact on the meaning of a term. "Intoxication," for example, has a potentially different meaning in a legal context than in a medical context. Some specific definitions correspond to operationalized designations developed for specific research projects and publications. However, much of the terminology currently used in the field of substance use problems and addictive behaviors is rooted in a medical–psychiatric framework and is strongly linked to diagnosis and taxonomy. Some terms are taken directly from medical terminology, for example, remission and relapse. Others are applications or extensions of terms in formal diagnostic taxonomies. The major psychiatric classification systems currently in use are the Diagnostic and Statistical Manual of Mental Disorders (DSM) produced by the American Psychiatric Association (currently in its 4th edition with the 5th edition released as of May 2013) and the generally similar International Classification of Diseases (ICD) produced by the World Health Organization (currently in its 10th edition with the 11th edition in development). DSM-IV substance use disorders are differentiated into categories of abuse (defined in regard to consequences) and dependence (as indicated by tolerance, withdrawal, and compulsivity). Other addictive behaviors are either classified separately (such as "pathological gambling," which is categorized as a disorder of impulse) or not specifically included.

DSM-5 combines the diagnoses of abuse and dependence into single substance designated disorders of graded clinical severity, adds a criterion of craving, and eliminates the abuse criterion of recurrent alcohol use-related legal problems. The reformulated diagnostic category now includes both substance use disorders

and non-substance addictions such as Gambling Disorder. From one perspective, the practical implications of the eevision of DSM-IV's classification of Substance-Related Disorders into DSM-5's classification of Substance Use and Addictive Disorders will be relatively minor in that all but one of the diagnostic criteria of DSM-IV has been carried over and only one new though related criterion has been added. Most individuals diagnosed with a substance use disorder under DSM-IV will also be diagnosed with a substance use disorder using DSM-5. Alternatively, the taxonomically modest combination of abuse and dependence may have major influence on the conceptualization of substance use disorders because the differentiating distinction of neuroadaptive dependence and disorder defined by harmful consequences resulting from persistent use is no longer reified by diagnostic categorization but is discouraged by the new nosology and terminology.

The classifications and terminologies of the DSM and ICD taxonomies serve a critical function for alcohol and drug use disorder diagnosis. They provide authoritatively endorsed operationalized criteria for psychiatrically identifying substance abuse and dependence syndromes according to articulated rules. These taxonomies are widely used and relied on, and are the practical standard for clinical and research purposes. Addictive behavior terminologies are so strongly anchored in the taxonomies that for the most part, their default meanings in professional contexts are the diagnostic classification distinctions and criteria.

Several issues should be noted. The DSM and ICD taxonomies are developed by experts in the mental health field and as scientific thought and empirical findings have changed over time, so have the concepts, categories, and criteria of the systems. Some degree of subjectivity, compromise, and supposition is inevitable. For example, it has been reported that the American Psychiatric Association DSM-III-R committee chose the term "dependence" over the term "addiction" by a single vote. The term addiction is becoming more established as a primary term and although it is not included in DSM-IV, the term "addictive disorders" is the category descriptor in DSM-5.

In developing and revising the taxonomies, judgments become formalized into established standards. Some are more structural, such as the number of symptoms required to meet diagnosis for dependence, while others reflect beliefs and conceptualizations, such as the hierarchical distinction of abuse and dependence in DSM-IV. Once terms and criteria have become standardized, they reify the concepts behind them. Clinicians typically adapt to conceptualizing patients with the current approved diagnostic system and researchers usually design projects that are at least compatible with the accepted taxonomy. Following DSM-IV criteria, the seeming actuality of dependence involving multiple symptoms and the hierarchical differentiation of the consequences implicit in abuse from the physiologically based reliance implicit in dependence is reinforced. The nosological framework of DSM substance use disorders, particularly the differentiation of dependence and abuse, is removed in DSM-5 and is likely to lead to changes in conceptualizations, perceptions, and perhaps even research findings about substance use symptoms and disorders.

Overgeneralization is a potential problem in the application of diagnostic terms. DSM and ICD criteria are intended primarily for case identification of psychopathology. The set of criteria for a particular disorder are somewhat like a screening instrument in that the items are intended to identify probable cases of a disorder including differentiation of similar or overlapping disorders. They are not necessarily the equivalent of a full description of the disorder. Relevant criteria may not be included for a variety of reasons including that the symptom may be common to multiple disorders, it may have a high co-occurrence with another criterion of the disorder, it may apply only to a small percentage of individuals with the disorder, it may be difficult to assess, and so on. An assumption that a term anchored in a diagnostic taxonomy is fully representative of the disorder may be inaccurate and misleading. Similarly, variables that are operationalized for the purpose of a research study are not necessarily a comprehensive delimitation of the term that they use to define. For example, a research study may define smoking a pack or more of cigarettes per day as an operationalization of nicotine dependence and the variable be interpreted by the researchers as representing nicotine dependence but it is not the full equivalent of nicotine dependence.

While delineating the fundamental nature and primary characteristics of substance use disorders is critical and controversial, determining the relationship of the symptoms of a disorder (the observable criteria) to the theorized construct of the disorder and articulating the relationship of the symptoms to the associated diagnostic terms and categories also raises significant issues. For example, is the implication of assigning a diagnostic term to a group of individuals: 1) that it designates those with a common set of observable symptoms without specification of the underlying causes or processes, 2) that it usefully corresponds to a group of individuals with related but diverse disorders, or, 3) that it identifies those who are assumed to have a specific underlying characteristic or dysfunction even if their manifest symptoms vary widely?

Assuming that diagnoses are made following the rules and criteria of taxonomy, many fundamental questions

typically remain unaddressed. For example, is it implied that a given diagnostic term has the same meaning when applied to everyone who meets the criteria for that term? Does the DSM-IV diagnostic term "opiate dependent" identify a homogeneous group? What is the relationship of those who meet one particular subset of the diagnostic criteria as opposed to those who meet a different subset, or to those who meet only some of the criteria (a subclinical subset) or even to those who meet none at all? Nevertheless, even when it is not clear exactly what a diagnostic term implies about the relationship of the observable symptoms to a theorized underlying pathology, use of the term in classifying individuals tends to reify belief in the reality of a discrete comprehensively conceptualized disorder.

Consideration of diagnostically standardized terms provides other examples of terminology both reflecting and shaping conceptualization. For the most part, 'craving' has been a central concept in the field. It emphasizes the experience of the alcohol- or drug-addicted individual and focuses on intense desire and impaired impulse control as important explanatory factors of addiction. To some it has a pejorative overtone. The term craving has not been included as a criterion in the DSM-IV or ICD-10 taxonomies for substance use disorders but apparently not because the term or the general concept is rejected. ICD-10 uses the term craving in the context of eating disorders but refers to strong desire or compulsion for substance use disorders. DSM-IV refers to persistent desire as a criterion for substance dependence even though its text notes that "Although not specifically listed as a criterion item, 'craving' (a strong subjective drive to use the substance) is likely to be experienced by most (if not all) individuals with Substance Dependence." The term craving may have been omitted from the criteria because it conveys stronger connotations of subjective experience and impulse control issues or perhaps for some incidental or procedural reason. However, once a lexicon is established it also reifies a particular emphasis. If clinicians and researchers do not ask about craving, they are less likely to be aware of the role of the phenomenon of craving to whatever extent it is different from strong desire, persistent desire, or compulsion. The differences are subtle but not meaningless. DSM-5 specifically incorporates craving as a criterion for substance use and addictive disorders.

Perhaps more importantly, the terms craving and compulsion refer to primary concepts describing presumably critical subjective experiences in addiction. They are the main experiential terms related to explanatory models of addictive behavior. Nevertheless, craving and compulsion are difficult to define, measure, or operationalize and are little more than synonyms for the phrase "strong desire." The new DSM-5 diagnostic criterion is worded as "Craving or a strong desire or urge to use…" Some terminology issues relate to disagreements about theory, pragmatic approaches, and research findings but others may be more a question of there simply not being a clear idea of what the concept is, what the term refers to, or how to measure it.

There are other issues involving the relationship of substance use problem terms and the concepts that they both reflect and shape. Problematic substance use nomenclature tends to include categorical rather than dimensional terminology and addictive disorder models frequently rely on dichotomous characterizations. For example, diagnoses do not currently allow for ratings of degree beyond the number of criteria met, relapse and abstinence are viewed as all or nothing, and quantity and frequency of the use of an illegal drug are often not even assessed. Many current addictive disorder models and the associated terminology promote a relatively homogenous perspective that minimizes the significance of subtypes. There is a strong emphasis on similarities and convergences of drug effects, neurological mechanisms, states of dependence, manifestations of abuse, and so on. Models of comorbidity are not well articulated and neither is the vocabulary designating co-occurrence of problem behaviors. For example, there is relatively little differentiation of single versus polydrug problematic substance use, there is no consensual lexicon to designate the range of possible relationships of co-occurring disorders, and there is no accepted terminology to describe the relationship of an event at one micro–macro level with an event at another level. This last example illustrates the poverty of language and conceptual frameworks available to discuss the relationship of a neurochemical depiction of addictive disorder with a behavioral depiction, to consider relationships of addictive behaviors with broader categories such as antisocial behavior, or to characterize variations in the relationship of predisposing factors to an outcome including whether they are earlier manifestations of that outcome or share a common underlying causal factor. Lastly, current addictive behavior terminology is not oriented toward developmental concepts, models involving transactional influences, or multidimensional factors. Without the words to describe particular types of factors and relationships, it is much harder to conceptualize and explore them.

RISK

Some of the terminology issues in the field of addictive behaviors relate to concerns that apply to most

fields. Issues related to risk are of concern across the range of behavioral and medical fields including the field of addictive behavior. The concept of risk is central to understanding substance use problem etiology, to early identification of vulnerable populations, and to effective prevention. Risk addresses questions regarding the relationship between observable antecedent and/or co-occurring factors and adverse outcomes.

At a basic level, the term "risk" asks, "given the presence of factor (or factors) X, what is the probability that outcome Y will occur?" For example, what is the likelihood that adolescents who have friends who use illegal drugs will themselves use drugs? In this sense, risk is a statement of predictability. In a typical approach, epidemiological data are used to identify factors that are statistically associated with, for example, substance use disorders, and then these factors are used as predictors of substance use disorders. Individuals in that particular data set who have these characteristics, or "risk factors," have a greater statistical probability of also currently or subsequently having the associated alcohol or drug use disorder outcome. Assuming that the epidemiological sample is representative, it is assumed that the predictive risk factors generalize to other populations.

However, whether the original epidemiologically derived risk factors were antecedent or concurrent with the substance use disorder outcome, the association is correlational. As such, while the identified risk factors may be useful markers, signals, or even early identifiers of substance use disorders, they are not demonstrated causes of or even proven contributors to substance use disorders.

Although this is widely recognized in all scientific fields, the use of risk as a probabilistic statement seems to be taken by some to have implications for causality. This is not surprising as there is ambiguity in the use of the term and risk has both a statistical and a causal meaning. In common parlance, risk often refers to circumstances that presumably influence events in a causal way, for example, stating that a person risks having a car accident if they drive while intoxicated or stating that individuals risk progressing to heroin dependence if they use marijuana. Even if the original intent of the risk statement was to report an observation of an association and does not intentionally imply causality, care is required to avoid an inference of causality on the part of the reader of the risk statement.

Analyzing conceptual and practical issues in psychiatric research, Helena Kraemer and colleagues have described how different statements about risk–outcome relationships may use a similar terminology but may be referring to significantly different types of associations. They examined different meanings underlying risk vocabulary and proposed a standard terminology to be used in discussions about risk. Although all of the variations in Kraemer et al.'s risk typology include at least a correlation between factor and outcome and therefore they contribute to the prediction of the outcome given the presence of the factor, they vary in the extent to which there is a presumed or demonstrated causal association. For risk–condition relationships where both correlation and temporal order have been established, they propose use of the following terms:

- *correlate*: a correlated factor; although each correlate may contribute to prediction of the other, neither of the correlated factors is known to be antecedent or causally determinative of the other;
- *risk factor*: referring only to an antecedent correlated factor;
- *marker*: an antecedent correlated factor with no presumed causal involvement in the outcome;
- *causal risk factor*: an antecedent correlated factor that presumably does causally influence the outcome.

Despite the cogency and importance of the points and proposals made by Kraemer and colleagues, the distinctions of risk terminology are often not observed and this leads to problems of miscommunications and invalid inferences. The issues are not scientific disputes about inferring causality from correlations. It is clear that even when it has been established that there is both an antecedent and a correlational relationship between a factor and an outcome, this does not demonstrate a causal relationship even if a plausible mechanism for causal influence can be hypothesized. There are other possibilities including that the antecedent may be a proxy for another influence, there may be an underlying common cause for the factor and outcome (third-factor alternatives), the antecedent may be an earlier form or stage of the outcome, and so on. Unless there is evidence to the contrary, it is safest to assume that markers are no more than correlates, not demonstrated causal influences and therefore not causal risk factors.

A significant challenge for a proposal of any standardized terminology is that unless it is sponsored by a major authoritative organization, it is often not adopted as a standard and if it uses idiosyncratic terms or definitions it is less likely to be used or used correctly. Formulating a set of terms that are explicit as to both denotation and connotation and that avoids relying on words that have multiple common definitions may help with both clarity and utility. Distinctions between causal influences and markers might be more easily and more often applied if the terminology were simpler and more clearly distinguished types of associated factors. Therefore, it would be helpful to reserve any use of the phrase "risk factors" to only include Kraemer's category of causal risk factors and to refer

to any other correlated factor as a "marker" independent of whether there is presumed antecedence. If a temporal distinction is needed, markers determined to exist prior to the predicted outcome could be designated as "antecedent markers." Using the term "risk factor" in this more limited definition may have the added advantage of not inadvertently implying causality for a variable that is only a proxy for the causal factor. Given the scientific complexity of the concept of causality, substituting the concept of having an influence on the occurrence of an outcome and using the term "predisposing" may also be more useful. Following is a revision of Kraemer's proposal. For risk–condition relationships where both correlations have been established, the following terms are suggested:

- *correlate*: one of two (or more) correlated factors; while each of the correlated factors may contribute to prediction of the other, neither of the correlated factors is identified as the outcome or known to necessarily be antecedent or causally determinative of the other, although such a relationship is not excluded as being possible;
- *marker*: a correlated factor with no implication of causal involvement in the outcome; markers are often more readily observable factors in a particular study or circumstance;
- *antecedent marker*: a correlated factor that is shown to be antecedent to the outcome but with no implication of causal involvement in the outcome;
- *predisposing factor*: an antecedent correlated factor that presumably does causally influence the outcome.

The intention of this revision of Kraemer et al.'s risk terminology is to propose a set of terms with the clearest possible definition and the most explicit implications. However, to be as useful as possible, the lexicon must correspond to common scientific thinking about the issue of risk. Contradicting, diverging from, or simply not corresponding to generally held conceptualizations often results in a terminology being misused or not adopted.

Implicit in viewing substance abuse and dependence as pathologies or medical disorders is the assumption that substance use problems are neither healthy nor normal. Therefore, there must be factors that cause the abnormal disease conditions. Construing problematic substance use as a psychopathology implies that it is involuntary, that it is an ailment that happens to individuals, and that there are likely to be physiological, psychological, and environmental contributors to its occurrence and course. Not everyone becomes a substance abuser, not even everyone who has used often abused substances. Therefore, there must be some factors that account for individual differences and increase the likelihood and severity of the disorder for some. Similarly, there are presumably factors that reduce the likelihood and severity for others. Unless and until they change, these concepts are at the core of our basic thinking about problematic substance use and any useful terminology must be correspondent with these fundamental assumptions.

However, it is a narrow linguistic distinction between stating that particular factors are antecedent and have a strong correlational association with an outcome as opposed to stating that particular factors influence that outcome. To reflect strongly entrenched common concepts of risk and to reinforce the distinction of correlated from predisposing factors, a comprehensive terminology of risk must include terms that identify predisposing factors that are thought to be predisposing or not as related to desirable versus undesirable outcomes.

Further expanding the revision of Kraemer's terminology, predisposing factors would be subtyped as follows:

- *beneficial predisposing factor:* a predisposing factor associated with a desirable outcome;
- *detrimental predisposing factor:* a predisposing factor associated with an undesirable outcome;
- *protective predisposing factor:* a predisposing factor associated with the nonoccurrence of an undesirable outcome.

Given that a correlation does not support inference about the operation or mechanism of the association, the determination of whether an antecedent marker is correlated with a desirable or undesirable outcome is a matter of conjecture. Without some assertion of cause, it is not possible, for example, to differentiate whether an antecedent marker has an association with the nonoccurrence of an event or with the occurrence of a different event.

Illustrating this, if parental monitoring is found to be an antecedent marker of lower levels of drug use in adolescents, it may be that monitoring has a direct positive influence on child behavior. Alternatively, it may be the case that monitoring is a signal that a range of parent–child circumstances are relevant and perhaps, for example, parents who monitor their children's activities also are better, more influential parents in many ways including having taught values associated with not using illicit drugs. Given the limitations of correlational information, the only judicious distinction would be to add a statistical description of the correlation and refer to either a positively or negatively correlated antecedent marker of an outcome.

While the concept of risk is often applied to subgroups (e.g. males) or those having some degree of a characteristic (e.g. impulsiveness), risk is often attributed to individuals as a personal characteristic in which case it is usually referred to as vulnerability or resilience. If a risk level label simply reflects that an individual has

a quality found for a group to be a marker or predisposing factor for an outcome, then the terms are not misleading particularly if it is recognized that each member of a group does not necessarily have the same level of risk attributable to the group as a whole. Alternatively, if the risk label is taken to mean that there is a specific quality of vulnerability or resilience and that these qualities are predisposing characteristics, then this is questionable unless there is evidence of such a general independent trait.

Attention to terminology may seem to be an unexciting procedural distraction from important issues. However, terminology is the language of scientific thought and communication and these both represent and shape understanding. Inconsistent or muddled terminology usually reflects concepts and models that are inadequate and counterproductive. Addictive behavior terminology is as controversial as other issues in the field. If that were not the case, this chapter would have been no more than a simple glossary.

DISCLAIMER

The views and opinions expressed in this chapter are those of the author and should not be construed to represent the views of the National Institute on Drug Abuse (NIDA), the National Institutes of Health (NIH), or any of the sponsoring organizations, agencies, or the US government.

Glossary

Diagnostic problematic substance use terms authoritatively endorsed operationalized sets of criteria designating nosological categories for psychiatrically identifying substance abuse and dependence syndromes according to articulated rules developed primarily for case identification of substance use-related psychopathology.

Problematic substance use maladaptive behavior defined by persistent and/or uncontrolled proscribed use of pharmacologically psychoactive substance(s) despite hazard or harm.

Risk in scientific contexts, risk is a probabilistic statement of the relationship between observable antecedent and/or co-occurring factors and adverse outcomes. Although useful in predicting outcomes, risk statements are usually based on correlational observations and when this is the case, risk does not imply causality.

Scientific terms formal, systematic definitions that provide a precise, standardized, documented nomenclature ensuring that statements about scientific information, observations, and theories are explicit in meaning, concrete in applicability, and replicable in the sense that the terms will have the same interpretations and specifiable implications for all users.

Further Reading

American Psychiatric Association, 2000. Diagnostic and Statistical Manual of Mental Disorders, Text Revision, fourth ed. American Psychiatric Association, Washington, DC.

Babor, T.F., Hall, W., 2007. Standardizing terminology in addiction science: to achieve the impossible dream. Addiction 102 (7), 1015–1018.

Borsboom, D., 2008. Psychometric perspectives on diagnostic systems. Journal of Clinical Psychology 64 (9), 1089–1108.

Courtwright, D.T., 2005. Mr. ATOD's wild ride: what do alcohol, tobacco, and other drugs have in common? The Social History of Alcohol and Drugs 20 (1), 105–140.

Glantz, M.D., Conway, K.P., Colliver, J.D., 2005. Drug abuse heterogeneity and the search for subtypes. In: Sloboda, Z. (Ed.), Drug Abuse Epidemiology. Springer Publishing, New York, pp. 15–27.

Glantz, M.D., Sloboda, Z., 1999. Analysis and reconceptualization of resilience. In: Glantz, M.D., Johnson, J.L. (Eds.), Resilience and Development: Positive Life Adaptations. Kluwer Academic/Plenum Press, New York, pp. 109–126.

Glantz, M.D., 2010. Touchstone issues for theories of substance abuse-dependence etiology. In: Scheier, L.M. (Ed.), The Handbook of Drug Use Etiology. American Psychological Association Press, Washington, DC, pp. 51–69.

Kraemer, H.C., Kazdin, A.E., Offord, D.R.M., Kessler, R.C., Jensen, P.S., Kupfer, D.J., 1997. Coming to terms with the terms of risk. Archives of General Psychiatry 54 (4), 337–343.

Leshner, A.I., 1997. Frontiers in neuroscience: the science of substance abuse: addiction is a brain disease, and it matters. Science 278 (5335), 45–47.

Maj, M., 2005. 'Psychiatric comorbidity': an artifact of current diagnostic systems? The British Journal of Psychiatry 186, 182–184.

Meehl, P.E., 2001. Comorbidity and taxometrics. Clinical Psychology: Science & Practice Winter 8 (4), 507–519.

O'Brien, C., Volkow, N., Li, T.K., 2006. What's in a word? Addiction versus dependence in DSM-V. American Journal of Psychiatry 163 (5), 764–765.

United Nations International Drug Control Programme, 2000. Demand Reduction: A Glossary of Terms (ODCCP Studies on Drugs and Crime: Guidelines). United Nations, New York.

Whorf, B.L., 1956. Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf. MIT Press, Cambridge, MA:

Widiger, T.A., Clark, L.A., 2000. Toward DSM-V and the classification of psychopathology. Psychological Bulletin 126 (6), 946–963.

World Health Organization, 1992. The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines. World Health Organization, Geneva, Switzerland:

Relevant Website

World Health Organization, 2011. Lexicon of Alcohol and Drug Terms. Retrieved from WHO 07/04/2011. http://www.who.int/substance_abuse/terminology/who_lexicon/en/.

CHAPTER

3

Epidemiology of Addiction

*Miriam C. Fenton**, *Christina Aivadyan*[$], *Deborah Hasin*[**]

[*]Columbia University, Mailman School of Public Health, New York State Psychiatric Institute, NYSPI/Columbia University, New York, NY, USA [$]New York State Psychiatric Institute, NYSPI/Columbia University, New York, NY, USA [**]Columbia University, College of Physicians and Surgeons, Mailman School of Public Health, New York State Psychiatric Institute, NYSPI/Columbia University, New York, NY, USA

OUTLINE

INTRODUCTION

The study of addiction is a multidisciplinary endeavor, encompassing a diverse range of fields such as biology, psychology, internal medicine, psychiatry, genetics, and epidemiology. In this chapter, we will present the epidemiological approach to addiction and provide an overview of our current knowledge of addiction epidemiology in the United States.

DEFINITION OF ADDICTION

Conceptually, addiction can be defined as "a primary, chronic, neurobiologic disease, with genetic, psychosocial, and environmental factors influencing its development and manifestations" characterized by impaired control or compulsive engagement in a specified behavior despite knowledge of harmful consequences. In the United States, addiction is operationally defined by

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00003-6

TABLE 3.1 DSM-IV-TR Substance Use Disorders

Diagnostic variables	Abuse	Dependence
Diagnostic criteria		
Recurrent substance use resulting in a failure to fulfill major role obligations	X	–
Recurrent use in situations in which it is physically hazardous	X	–
Recurrent substance-related legal problems	X	–
Continued use despite persistent or recurrent substance-related problems	X	–
Tolerance	–	X
Withdrawal*	–	X
Persistent desire or unsuccessful efforts to cut down or control use	–	X
Using in larger amounts or over a longer period than the person intended	–	X
Neglect of important activities because of substance use	–	X
A great deal of time spent in substance-related activities	–	X
Continued use despite substance-related physical/psychological problems	–	X
Diagnostic threshold	1+ criteria	3+ criteria

* *Marijuana and hallucinogen use disorders do not include the withdrawal criterion.*

the *Diagnostic and Statistical Manual* (DSM) as substance abuse or dependence (substance use disorder refers to a diagnosis of abuse and/or dependence). Diagnostic criteria for substance abuse and dependence from the most recent version of the DSM, the DSM-IV-TR, are provided in Table 3.1. While other addictive behaviors exist and are discussed throughout the *Encyclopedia of Addictive Behaviors*, the present chapter will only consider DSM-defined abuse and dependence on substances which produce intoxication, including alcohol, illicit drugs, and prescription drugs. In order to provide a context for the discussion of substance use disorders, we will also briefly consider substance use. However, it is important to note that substance use is not necessarily indicative of addiction and not all substance users develop a substance use disorder.

DEFINITION OF EPIDEMIOLOGY

The field of epidemiology has two primary objectives. First, epidemiological investigation aims to obtain information on the distribution of health outcomes within a population. Distribution includes prevalence, incidence, and persistence. Prevalence is the number of cases of a health outcome that exist within a given population at a specific point in time. Incidence refers to the number of new cases of a health outcome that arise in a given population over a certain time period whereas persistence refers to the proportion of cases who continue to experience a health outcome over a certain time period. Second, epidemiology is concerned with the determinants of a health outcome within a given population, which refers to risk factors for the onset and persistence of health outcome.

Accordingly, this chapter will present information on the distribution and determinants of substance use, abuse and dependence in the United States. In part 1, we discuss the epidemiology of substance use in the United States and provide an overview of the prevalence of alcohol and drug use. In part 2, we present information on the epidemiology of substance use disorders in America. We begin this section with a description of the key substance use disorder epidemiology studies conducted in the United States since 1980. Next, we provide findings from these studies on the distribution of alcohol and drug use disorders including prevalence, incidence, and persistence. Lastly, we will consider epidemiological findings on the determinants of substance use disorders including sex, race, and age.

PART 1. THE EPIDEMIOLOGY OF SUBSTANCE USE IN THE UNITED STATES

Alcohol consumption can be recorded at the population level and individual level, while drug consumption is primarily measured at the individual level.

To determine population-level alcohol consumption, "apparent per capita consumption" is calculated by dividing the total amount of alcohol consumed in the United States given records of alcohol sales by the total US population aged 14 and older.

Alcohol and drug use can be measured at the individual level in United States surveys with questions about the quantity and frequency of consumption. In this chapter we will include data on self-reported alcohol and drug use obtained by the following two important population-level surveys:

1. The National Survey on Drug Use and Health (NSDUH): Funded by the Substance Abuse and Mental Health Services Administration, the NSDUH is a large annual national survey of ages 12 and older designed to track trends in substance use and collects data on current use via self-administered computerized interview. The NSDUH is the most comprehensive and recent source on substance consumption among adolescents and adults in the United States.
2. Monitoring the Future (MTF): Designed to track substance use behaviors, attitudes, and values and funded by the National Institute of Drug Abuse (NIDA), annual MTF surveys are conducted among a large nationally representative cross-sectional sample of US adolescents. The MTF surveys have been conducted annually since 1975 and thus provide an important and unique opportunity to understand trends over time in substance use among adolescents in the United States.

Alcohol

Per Capita Consumption

Figure 3.1 demonstrates changes over time in per capita alcohol consumption in the United States. In 2007, the most recent data available, 2.31 gallons of alcohol were consumed per capita, compared to 2.27 gallons per capita in 2006 and 2.23 gallons per capita in 2005. Indeed, alcohol consumption appears to have steadily increased since 1995 when a low of 2.15 gallons per capita was reported. However, present alcohol consumption is lower than the high level recorded in the early 1980s (2.76 gallons per capita).

NSDUH

Table 3.2 provides information from the NSDUH on individual alcohol use among US teenagers and adults in 2005 and 2006 in three time frames: past month, past year, and lifetime. Findings for 2005 and 2006 are similar. In the total population, approximately 83% report any alcohol consumption in their lifetime, around two thirds (approximately 66%) report past year consumption and about half report past month alcohol use.

The NSDUH sample consists of individual aged 12 and older. However, since alcohol sales under the age of 21 are illegal in the United States, it is important to

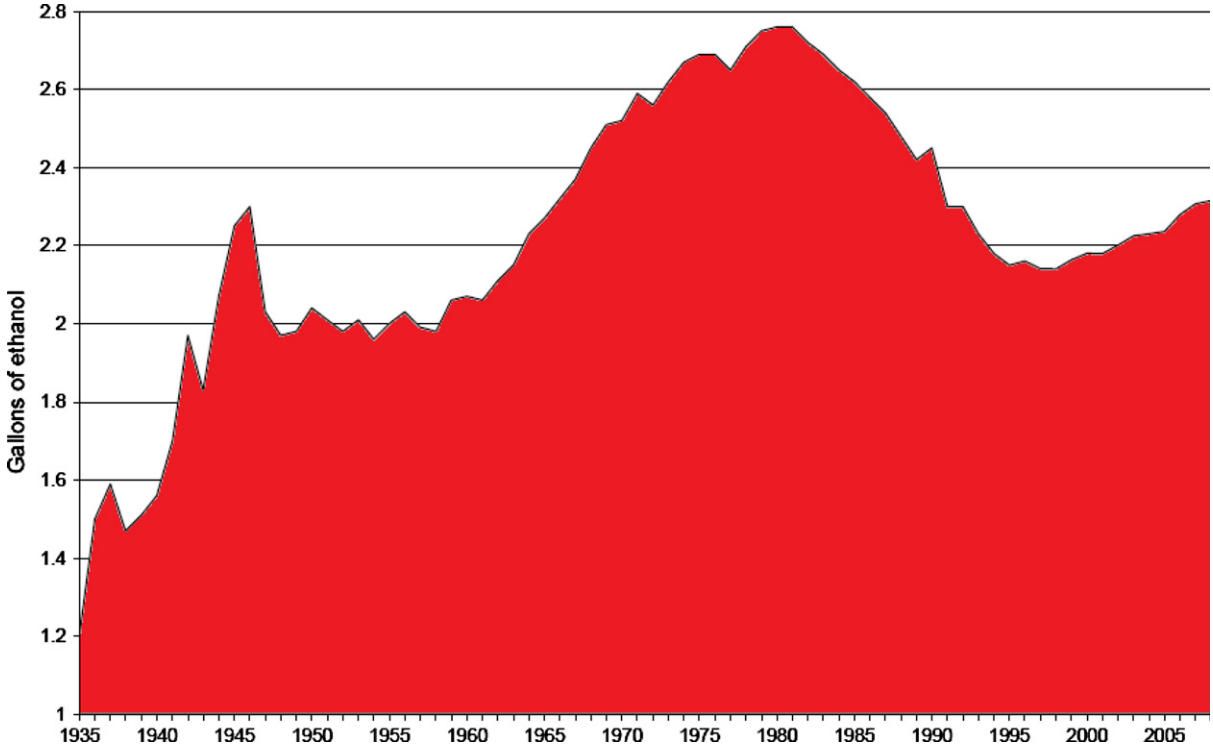


FIGURE 3.1 US per capita alcohol consumption. LaVallee, R., Yi, H. *Surveillance Report #90: Apparent Per Capita Alcohol Consumption: National, State, and Regional Trends, 1977–2008.* Bethesda, MD: NIAAA (National Institute on Alcohol Abuse and Alcoholism), Division of Epidemiology and Prevention Research, Alcohol Epidemiologic Data System, August 2010.

TABLE 3.2 Alcohol Use in the NSDUH

	Prevalence (%)					
	Past month		Past year		Lifetime	
	2005	2006	2005	2006	2005	2006
Total	51.8	50.9	66.5	66.0	82.9	82.7
Ages						
12–17	16.5	16.6	33.3	32.9	40.6	40.4
18–25	60.9	61.9	77.9	78.8	85.7	86.5
26+	55.1	53.7	69.0	68.3	88.2	87.7

consider how this measure varies by age. Less than half (around 40%) of individuals aged 12–17 (for whom purchasing alcohol is illegal in the United States), report any lifetime alcohol use, whereas the prevalence of lifetime alcohol use among the older age groups is more than double that. Although lifetime alcohol use is similar for 18–25-year olds and 26+ year olds, recent consumption of alcohol use is highest among 18–25-year olds (see Table 3.2). For example, in 2006, 62% of 18–25-year olds report past month alcohol use compared to around 17% of 12–17-year olds and 54% of those aged 26 and older. However, while alcohol use is generally homogenous among individuals older than 26, a great deal of heterogeneity of alcohol consumption exists among individuals aged 12–25 years (see Fig. 3.2).

MTF

Monitoring the future surveys reveals that the prevalence of alcohol use among US adolescents declined steadily in the last two decades (see Fig. 3.3). For example, past month prevalence of alcohol use was 25% in 1991 and 15% in 2009 for 8th graders, 43% in 1991 and 30% in 2009 for 10th graders, and 54% in 1991 and 44% in 2009 for 12th graders.

Similar declines in the prevalence of binge drinking (defined as having five or more drinks in a row at least once in the prior 2 weeks) are evident for 8th (11% in 1991 compared to 8% in 2009), 10th (21% in 1991 compared to 18% in 2009), and 12th graders in the United States (30% in 1991 compared to 25% in 2009). However, as shown in Fig. 3.4, adolescent reductions in alcohol use may be reaching a plateau. Data obtained from MTF on past month, past year, and lifetime alcohol use by teenagers in 2007, 2008, and 2009 are presented in Fig. 3.4.

Drugs

NSDUH

Data from NSDUH on past year and lifetime drug use in 2005, 2006, and 2007 are presented in Table 3.3. Marijuana is the most commonly used drug with approximately 40% of the US population reporting lifetime use and approximately 10% reporting past year use. Nonmedical prescription drug use is second to marijuana in terms of prevalence. Of the prescription drug classes, nonmedical use of prescription painkillers

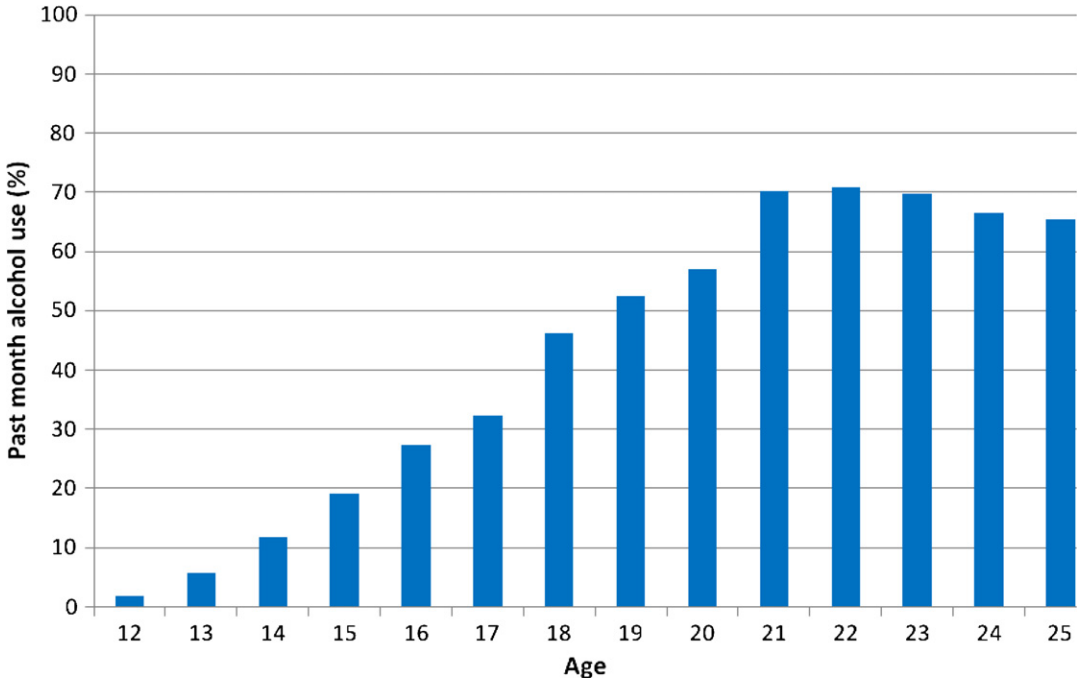


FIGURE 3.2 Past month alcohol use among ages 12–25.

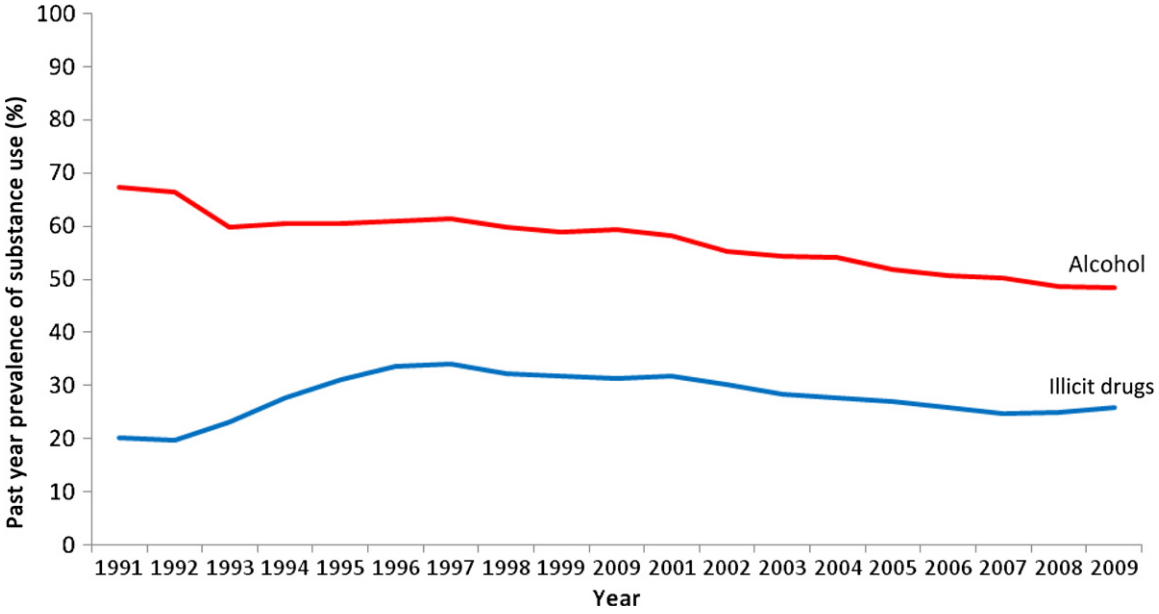


FIGURE 3.3 Past year prevalence of adolescent substance use from 1991 to 2009.

is most frequent, followed by stimulants, tranquillizers, and sedatives. Heroin was the least prevalent drug used, with only 0.1% of the population reporting past year use in 2007.

MTF

Prevalence of any illicit drug use among adolescents increased in the early 1990s and then began to decrease in 1995 for a decade, at which point it appears to start increasing (see Fig. 3.3 for past year illicit drug use). A more detailed summary picture of past month, past year, and lifetime prevalence of illicit drug use among 8th, 10th, and 12th graders in the United States is provided in Fig. 3.5. In general, prevalence of use between 2007 and 2009 was steady for 8th and 12th graders; however among 10th graders; use appeared to decrease between 2007 and 2008, and then increase between 2008 and 2009.

However, trends in drug use among American adolescents since the 1990s appear to differ somewhat by substance. Prevalence of marijuana use among teenagers decreased substantially in the 1990s and early 2000s. However, in recent years use appears to be increasing again. For example, past month prevalence of marijuana use among 8th graders was 11% in 1996, 6% in 2007, and 7% in 2009. Likewise, in 1996, 25% of 10th graders reported past month marijuana use, compared to 18% in 2007 and 19% in 2009, whereas among 12th graders this prevalence was 27% in 1997, 22% in 2007, and 24% in 2009. Ecstasy use peaked in 2001 followed by a sharp decrease until 2005 at which point use began to increase slightly and now appears to be reaching a plateau. Cocaine and heroin use increased in the 1990s, decreased in 2000, and presently appear to be relatively stable over time. Inhalant use has decreased since 1991 whereas nonmedical prescription painkiller, sedative, and tranquilizer use seems to be increasing.

PART 2. THE EPIDEMIOLOGY OF SUBSTANCE USE DISORDERS IN THE UNITED STATES

Substance Use Disorder Epidemiology Studies

Epidemiological investigation into substance use disorders in the United States began in the 1960s with community surveys of psychiatric impairment and subsequent functional impairment. However, these studies measured a general continuum psychopathology and therefore did not provide information on the distribution of specific disorders. Furthermore, these community studies assumed an entirely psychosocial causal mechanism of mental illness which has since been refuted by modern scientific developments in genetics, neuroscience, and psychopharmacology of psychiatric disorders that emerged later. As such, results presented in this chapter will be limited to findings obtained from epidemiologic studies conducted since the 1980s. In addition, to aid in interpretation, we will only include the five US-based epidemiological studies that adhered to the following design and methodological criteria:

1. Substance use disorder diagnoses were based on diagnostic nomenclatures of the American Psychiatric Association (APA) as presented in the

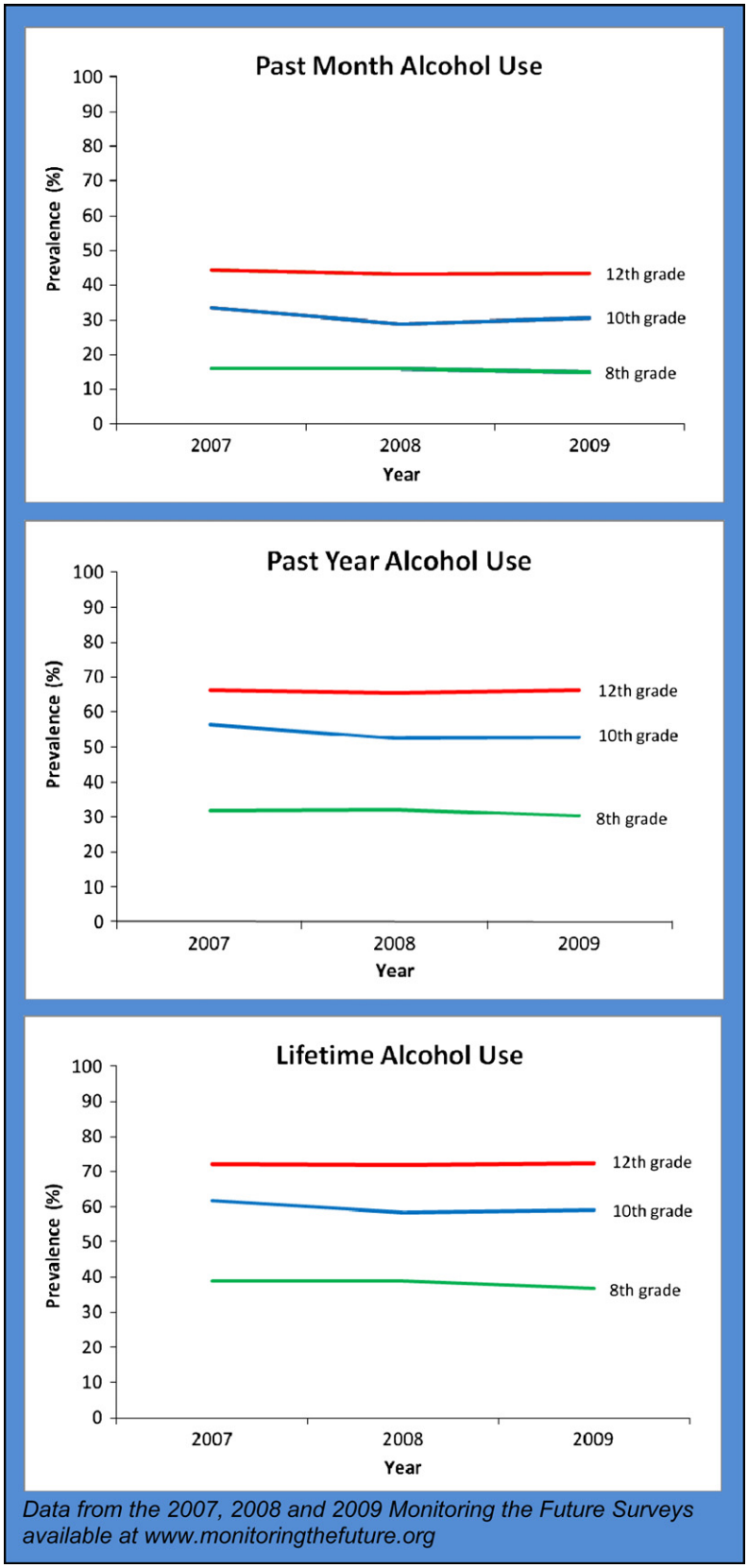


FIGURE 3.4 Past month, past year, and lifetime self-reported alcohol consumption from 2007 to 2009 among US adolescents.

Diagnostic and Statistical Manual of Mental Disorders, third edition (DSM-III), DSM-III-R (revised), or the DSM-IV obtained via standardized diagnostic interviews.
2. The study population was nationally or regionally representative, with a sample size (*N*) of at least 5000 that includes a broad range of adult ages.
3. Results include weighted lifetime and current estimates of distribution accompanied by an indicator of statistical precision such as confidence intervals or standard errors.
4. An explicit response rate was provided.

Epidemiological Catchment Area Study

The Catchment Area Study (ECA) was conducted between 1978 and 1983 in five US communities including New Haven, CT; Los Angeles, CA; Baltimore, MD; St. Louis, MO; and Durham, NC. The ECA obtained DSM-III substance use disorder diagnoses using the Diagnostic Interview Schedule (DIS), which was specifically designed for this purpose. The ECA was revolutionary in its sample size (20 000) and use of modern diagnostic criteria, and thus provided the first reliable and accurate weighted estimates of prevalence of mental health disorders in the United States. However, the ECA did not collect risk factor data in a standardized manner across all five sites, was geared toward treatment needs, and generalizability may have been limited by its use of only five sites.

TABLE 3.3 Prevalence of Drug Use in the NSDUH

	Prevalence (%)					
	Past year			Lifetime		
	2005	2006	2007	2005	2006	2007
Any drug	14.4	14.5	14.4	46.1	45.4	46.1
Cannabis	10.4	10.3	10.1	40.1	39.8	40.6
Cocaine	2.3	2.5	2.3	13.8	14.3	14.5
Heroin	0.2	0.2	0.1	1.5	1.5	1.5
Hallucinogens	1.6	1.6	1.5	13.9	14.3	13.8
Inhalants	0.9	0.9	0.8	9.4	9.3	9.1
Nonmedical prescription drug use	6.2	6.7	6.6	20.0	20.7	20.3
Prescription painkillers	4.9	5.1	5.0	13.4	13.6	13.3
Prescription stimulants	1.1	1.5	1.2	7.8	9.1	8.7
Prescription tranquilizers	2.2	2.1	2.1	8.7	8.7	8.2
Prescription sedatives	0.3	0.4	0.3	3.7	3.6	3.4

National Comorbidity Survey and National Comorbidity Study-Replication

The National Comorbidity Survey (NCS) was a psychiatric epidemiologic survey conducted between 1990 and 1992, with a sample of about 8000. The NCS collected risk factor data on a subset of the respondents. The assessment instrument used in the NCS was the University of Michigan version of the Composite International Diagnostic Interview (UM-CIDI), which obtained substance abuse and dependence diagnoses according to DSM-III-R criteria.

A decade later (2001–02), the National Comorbidity Study-Replication (NCS-R) was fielded, which aimed to obtain information on time trends in psychiatric disorders and provide estimates on the prevalence of psychiatric disorders according to DSM-IV criteria. Unfortunately, a major methodological problem with the NCS-R limits the conclusiveness and representativeness of alcohol and drug dependence results from this study. In brief, the version of the CIDI used in the NCS-R (the World Mental Health-CIDI (WMH-CIDI)) skipped respondents with no abuse symptoms from the dependence module. This is particularly concerning given later work indicating that 33.7% of individuals with current alcohol dependence do not meet criteria for alcohol abuse and 22.0% of individuals with current drug dependence do not meet criteria for drug abuse.

National Longitudinal Alcohol Epidemiologic Survey and National Epidemiologic Survey on Alcohol and Related Conditions

The National Longitudinal Alcohol Epidemiologic (NLAES) and National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) provide a wealth of information on the epidemiology of alcohol and drug use disorders and other psychiatric disorders due to their robust sampling and measurement methods and their unprecedented sample sizes. The NLAES was conducted in 1991–92 and included 42 862 respondents. Wave 1 of the NESARC was conducted in 2001–02 and included 43 093 respondents; 34 653 of whom were followed up 3 years later at Wave 2 (2004–05). The diagnostic interview used in the NLAES and NESARC was the Alcohol-Use Disorder and Associated Disabilities Interview Schedule–DSM-IV Version (AUDADIS-IV). This structured interview, designed for lay interviewers, was developed to advance measurement of substance use and mental disorders in large-scale surveys. Computer diagnostic programs implemented the DSM-IV criteria for the disorders using AUDADIS-IV data.

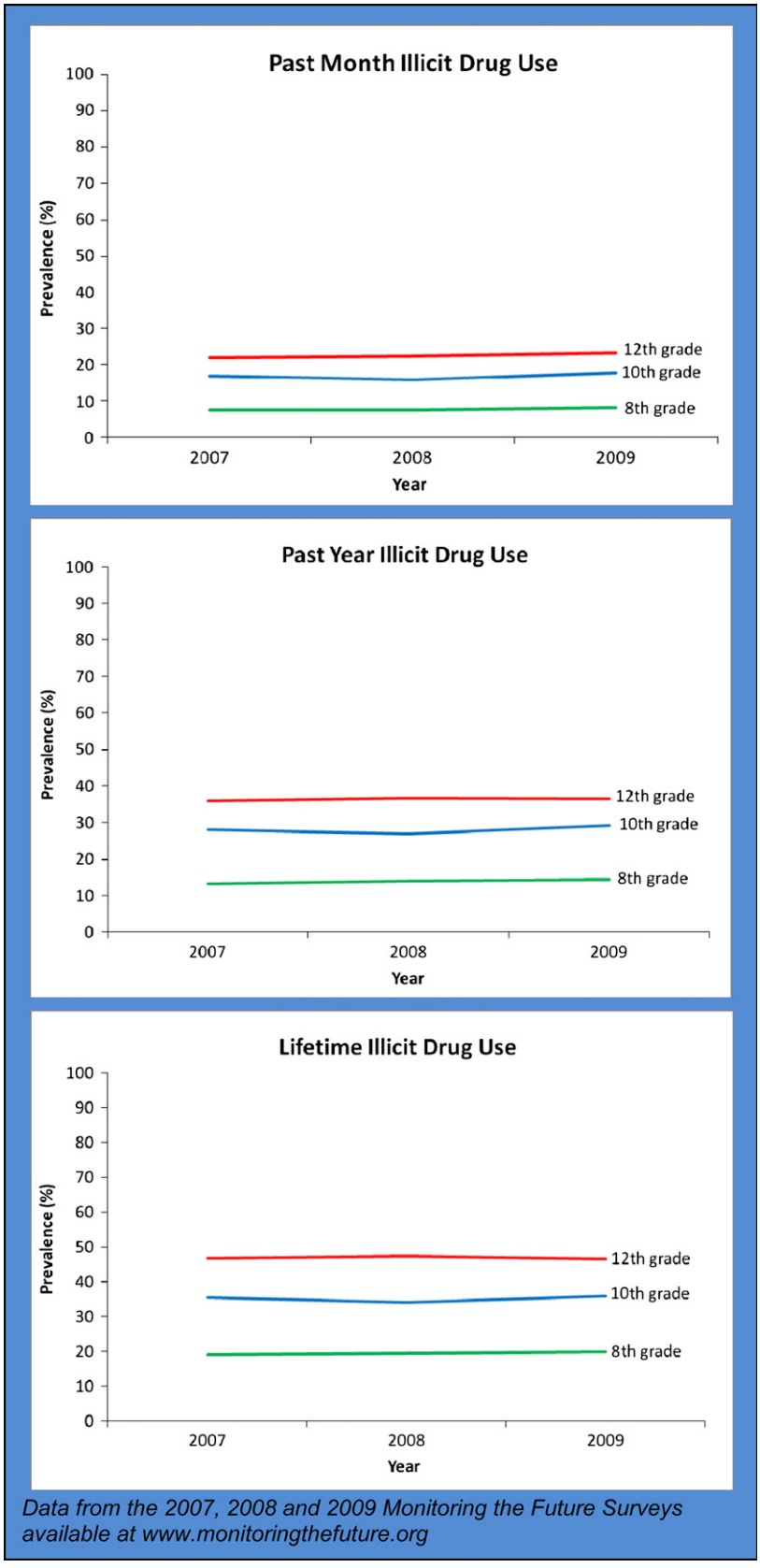


FIGURE 3.5 Past month, past year, and lifetime self-reported illicit drug use from 2007 to 2009 among US adolescents.

Extensive detailed AUDADIS-IV items operationalize DSM-IV abuse and dependence criteria for alcohol among all lifetime drinkers and for 10 drug classes (sedatives, tranquilizers, opiates (other than heroin), stimulants, hallucinogens, cannabis, cocaine (including crack cocaine), inhalants/solvents, heroin, and other drugs) among all lifetime drug users. Consistent with DSM-IV, lifetime diagnoses of alcohol or drug abuse required one or more of the four abuse criteria in the 12-month period preceding the interview or previously. AUDADIS-IV substance dependence diagnoses required three or more of the seven DSM-IV dependence criteria in the last 12 months or during any previous 12-month period. For prior diagnoses of dependence, three or more criteria must have occurred within a 1-year period following the DSM-IV clustering criterion. The withdrawal criterion of dependence diagnoses was measured as a substance-specific syndrome in accordance with DSM-IV.

AUDADIS-IV diagnoses of substance use disorders incorporate important improvements over other survey instruments, particularly the WMH-CIDI (used in the NCS-R), which skipped dependence questions among those with no abuse symptoms. By missing current dependence cases (especially likely among women and minorities), the WMH-CIDI underestimates dependence prevalence and limits the study of comorbidity between alcohol dependence and other psychiatric disorders. In contrast, the NESARC provides complete coverage of DSM-IV alcohol and drug dependence. The robustness of AUDADIS-IV substance use diagnoses is well documented, with good to excellent test–retest reliability, inter-rater reliability and convergent, discriminant and construct validity demonstrated in clinical and general population studies conducted in the United States and abroad.

Further information on the study design, sampling assessment methods of these five studies is provided in Table 3.4. While this chapter will provide findings from all five of the studies described above, given the methodological and measurement advantages of the NLAES and NESARC, we will primarily focus on results from these studies when discussing the epidemiology of alcohol and drug use disorders.

The Distribution of Substance Use Disorders in the United States

Prevalence of Alcohol-Use Disorders

Information on the prevalence of current and lifetime alcohol abuse, dependence, and any alcohol-use disorders using data obtained in the ECA, NCS, NLAES, NCS-R, and NESARC are presented in Table 3.5. The prevalence of current (last 12 months) alcohol abuse is lowest in the ECA (1.9%) and highest in the NESARC (4.7%), while current alcohol dependence was lowest in the NCS-R (1.3%) and highest in the NCS (7.2%). The lifetime prevalence of alcohol abuse ranges from 5.6% in the ECA to 17.8% in the NESARC, while alcohol dependence ranges from 5.4% in the NCS-R to 14.1% in the NCS. This range of alcohol abuse and dependence prevalence estimates is likely an artifact of between-study design and measurement discrepancies summarized in Table 3.5. For example, the very low prevalence of current alcohol dependence in the NCS-R may be explained by the skip pattern in the WMH-CIDI noted above.

Of the five studies listed in Table 3.4, the NLAES and NESARC have the most similar methodology, support for the reliability and validity of the diagnoses, and the largest sample sizes. Their results are less variable than other between-survey comparisons. As such, remaining alcohol-use disorder results presented in this chapter will focus on NLAES and NESARC findings.

Time trends: The high degree of comparability between the NLAES and the NESARC with regard to methodology and measurement of alcohol-use disorders enables analysis of trends in the prevalence of alcohol abuse and dependence in the decade between 1991–92 (when NLAES was conducted) and 2001–02 (when NESARC was conducted).

As shown in Table 3.5, the prevalence of alcohol dependence in the US population decreased significantly from 4.38% in 1991–92 to 3.81% in 2001–02 ($p < 0.01$). This decrease of 0.5% is consistent with other reports from national surveys indicating similar decreases in the percent of heavy drinking during this period. However, conversely the prevalence of alcohol abuse increased by 1.62% from 3.03% in 1991–92 and 4.65% in 2001–02, a statistically significantly change ($p < 0.01$). Grant and colleagues, the authors of this analysis, postulate decreases in the prevalence of heavy drinking and alcohol dependence could be associated with changes in drinking norms including more negative attitudes about maladaptive alcohol use. In turn, this may have led to an increase in alcohol-related social/interpersonal problems, which is a DSM-IV alcohol abuse criterion (see Table 3.1).

Prevalence of Drug Use Disorders

Lifetime and current prevalence of any drug use disorder in the ECA, NCS, NLAES, NCS-R, and NESARC is presented in Table 3.6. Prevalence of current drug abuse is similar in all five studies and ranges from 0.8% (NCS) to 1.4% (NCS-R and NESARC). However, current drug dependence is much higher in the NCS (2.8%) than in the ECA (0.9%), NESARC (0.6%), NLAES (0.5%), and NCS-R (0.4%). Furthermore, lifetime

TABLE 3.4 Study Design, Sampling Features, and Assessment Methods of the Five US Psychiatric Epidemiological Studies

		ECA	NCS	NLAES	NCS-R	NESARC
Study design	Sponsoring institution	NIMH	NIMH	NIAAA	NIMH	NIAAA + National Institute of Drug Abuse (NIDA)
	Years of data collection	1980–84	1990–92	1991–92	2001–03	2001–02
	Field work conducted by	Independent academic researchers at the five sites	Survey Research Institute, University of Michigan	US Bureau of the Census	Survey Research Institute, University of Michigan	US Bureau of the Census
	Follow-up component	1-year follow-up at all sites ($N = 10\ 167$), 13-year follow-up at the Baltimore site	10-year follow-up ($N = 4375$)	None	None	3-year follow-up ($N = 34\ 653$)
Sampling features	Sample size	20 219	8098	42 862	9282	43 098
	Response rate (approximate)	77.6%	82.6%	89.2%	70.9%	81.00%
	Sample	Five US communities	US population	US population	US population	US population
	Sampling method	Probability, block sampling and oversampling in some sites	Probability	Probability, oversampling for minorities and young adults	Probability	Probability, oversampling for minorities and young adults
	Individuals surveyed	Household + institutional residents	Household and college residents	Household residents	Household and college residents	Household and group quarters residents
	Age range	18 and older	15–54	18 and older	18 and older	18 and older
Diagnostic interview	DSM version	DSM-III	DSM-III-R	DSM-IV	DSM-IV	DSM-IV
	Diagnostic interview	DIS	UM-CIDI	AUDADIS	WMH-CIDI	AUDADIS-IV
	Time frame for "current"	Prior 6 months	Prior 6 months	Prior 12 months	Prior 12 months	Prior 12 months
	Diagnostic coverage	Substance, affective, anxiety, psychotic disorders	Substance, affective, anxiety, psychotic disorders	Substance, major depression	Substance, affective, anxiety, psychotic disorders	Substance, affective, anxiety, personality disorders, psychotic screening

TABLE 3.5 Weighted Prevalence of Current and Lifetime Alcohol Disorders in Five General Population Surveys

	Weighted prevalence of alcohol-use disorders (%)					
	Current*			Lifetime		
Epidemiology study (date)	Abuse	Dependence	Abuse/Dependence	Abuse	Dependence	Abuse/Dependence
ECA (1981–83)	1.9	2.8	4.8	5.6	7.9	13.5
NCS (1991–92)	2.5	7.2	9.7	9.4	14.1	23.5
NLAES (1991–92)	3	4.4	7.4	4.9	13.3	18.2
NCS-R (2001–02)	3.1	1.3*	4.4**	13.2	5.4	18.6*
NESARC Wave 1 (2001–02)	4.7	3.8	8.5	17.8	12.5	30.3
NESARC Wave 2 (2004–05)	4.9	4.1	9.7	19.3	14.5	34.6

** Current is defined as past 6 months in the ECA and past 12 months in the NCS, NLAES, NCSR, and NESARC (Waves 1 and 2).*
*** Dependence not assessed in those without abuse.*

TABLE 3.6 Weighted Prevalence of Current and Lifetime Drug Disorders in Five General Population Surveys

	Weighted prevalence of drug use disorders (%)					
	Current*			Lifetime		
Epidemiology study (date)	Abuse	Dependence	Abuse/Dependence	Abuse	Dependence	Abuse/Dependence
ECA (1981–83)	1.2	0.9	2.0	3.5	2.6	6.1
NCS (1991–92)	0.8	2.8	3.6	4.4	7.5	11.9
NLAES (1991–92)	1.1	0.5	1.5	3.1	2.9	6.1
NCS-R (2001–02)	1.4	0.4**	1.8	7.9	3.0*	10.9
NESARC Wave 1 (2001–02)	1.4	0.6	2.0	7.7	2.6	10.3
NESARC Wave 2 (2004–05)	1.7	0.8	2.4	10.2	3.4	12.0

** Current is defined as past 6 months in the ECA and past 12 months in the NCS, NLAES, NCSR, and NESARC (Waves 1 and 2).*
*** Dependence not assessed in those without abuse.*

prevalence of drug dependence is also substantially higher in the NCS.

At the time of the initial publication of NCS results, considerable attention was drawn to the substantially higher prevalence of many disorders in the NCS and drew a number of efforts to explain the findings. This is particularly pertinent for current and lifetime drug dependence, which is markedly different to all other findings. The explanations largely focused on differences between the diagnostic interview used in the NCS and the assessment procedures used in other surveys. For example, the drug dependence gateway questions were asked at the beginning of the CIDI to prevent individuals from learning that gateway questions lead to more extensive questioning. Thus, more respondents may have been screened into the drug disorder module in the NCS than in other studies which may have increased the prevalence of drug dependence in the NCS. In addition, while the ECA, NCS-R, NLAES, and NESARC include ages 18 and older, the NCS sample is limited to 15–54-year olds. Since prevalence of drug dependence decreases with age, exclusion of the oldest age groups in the NCS may have inflated the prevalence in this study compared to other studies. Given that prevalence of drug abuse in the NLAES, which was conducted around the same time as the NCS, is similar to the prevalence obtained in preceding (ECA) and following (NCS-R, NESARC) studies, measurement differences between the NCS and other surveys appear a reasonable explanation for the high rates found by the NCS.

Time trends: As with alcohol, we can discern time trends in the prevalence of drug abuse and dependence in the decade between 1991–92 (when NLAES was conducted) and 2001–02 (when NESARC was conducted), due to the high degree of comparability between these studies.

Compton and colleagues revealed that although the proportion of cannabis users remained steady between 1991–92 and 2001–02, the past year cannabis use disorder prevalence increased significantly from 1.2 to 1.5% in the population ($p = 0.01$), and from 30.2 to 35.6% among past year cannabis users ($p < 0.01$). Blanco and colleagues demonstrated that both the prevalence of

past year nonmedical prescription drug use and past year nonmedical prescription drug abuse/dependence increased significantly between 1991–92 and 2001–02 ($p < 0.001$). However, the increase from 19.9 to 23.6% of nonmedical prescription drug use disorders among past year nonmedical users was not significant ($p = 0.15$).

Incidence and Persistence

Data on incidence and persistence are required to understand the etiology and course of alcohol-use disorders. However, while prevalence estimates can be obtained from cross-sectional data, large, prospectively observed samples are required to generate accurate incidence and persistence estimates. Psychiatric epidemiological studies have rarely been able to provide longitudinal data because the costs and practical difficulties are so great. However, some information on incidence of substance abuse and dependence from the ECA and NESARC has been published and information on the persistence of substance dependence between Waves 1 and 2 of the NESARC is also available.

One-year incidence of alcohol abuse/dependence was 1.79% using combined information from four ECA sites (Baltimore, St. Louis, Durham, and Los Angeles). In the NESARC, 1-year incidence of alcohol abuse was 1.02% and 1-year incidence of alcohol dependence was 1.70%.

Although a popular clinical conceptualization of alcohol dependence is that of a chronic and relapsing disease, an analysis of the persistence of alcohol dependence between Waves 1 and 2 of the NESARC indicated that this disorder persisted in only 30.1% of respondents with alcohol dependence at baseline.

One-year incidence of drug abuse/dependence was 1.09% in the ECA using data from the Baltimore, St. Louis, Durham, and Los Angeles sites. In the NESARC, 1-year incidence of drug abuse was 0.28% and 1-year incidence of drug dependence was 0.32%.

As with alcohol dependence, clinicians often conceptualize drug use disorders as chronic. However, similar to the findings on alcohol dependence, in the NESARC, 3-year persistence of any drug use disorder was 30.1% and 3-year persistence of a cannabis use disorder was 31.5%.

Determinants (Risk Factors) of Substance Use Disorders in the United States

Alcohol

- Sex

  Epidemiological studies consistently indicate that males have an increased risk of developing an alcohol-use disorder compared to females. Indeed, the male to female ratio of lifetime alcohol dependence was 5.0:1 in the ECA, 2.5:1 in the NCS, and 2.2:1 in the NLAES. The prevalence of current alcohol abuse was 4.7% among males and 1.5% among females in the NLAES, and 6.9% among males and 2.6% among females in the NESARC. Similarly, the prevalence of current alcohol dependence was 6.3% among males and 2.6% among females in the NLAES, and 5.4% among males and 2.3% among females in the NESARC. NESARC prevalence estimates are shown in Fig.3. 6. Furthermore, males were significantly more likely than females to develop incident alcohol abuse (odds ratio (OR), 2.3) and dependence (OR, 2.4) between Waves 1 and 2 of the NESARC, controlling for Wave 1 demographic characteristics and psychiatric disorders ($p < 0.01$).

  Of note, clinicians often conceive of alcohol disorders among women as "telescoped" with a later onset of alcohol use but shorter times from use to dependence and treatment. However, in a recent analysis of NLAES and NESARC data, Keyes and colleagues found little evidence for a telescoping effect among women in the population. Furthermore, sex differences in the prevalence of alcohol-use disorders appear to have decreased over time. As a result, younger women may require more targeted prevention and intervention efforts.
- Race/ethnicity

  Alcohol-use disorders are most prevalent among Native Americans and least prevalent among Blacks and Asians. This is true for alcohol abuse and dependence. For example, in the NLAES, among Native Americans 8.1% met criteria for alcohol abuse and 9.0% met criteria for alcohol dependence, while among Asians 1.1% met criteria for alcohol abuse and 2.3% met criteria for alcohol dependence. Figure 3.6 presents information on the prevalence of alcohol abuse and dependence stratified by five racial groups. Compared to Whites, the odds of current and lifetime alcohol abuse and dependence were lower among Blacks and Asians and higher among Native Americans ($p < 0.05$).

  Despite this, incident alcohol disorders in the NESARC varied little by race (note: this analysis did not include Native Americans, possibly due to small sample size). Indeed only one significant race difference was found – Blacks were significantly less likely than Whites and Hispanics to report incident alcohol abuse (OR, 0.6) at Wave 2 of the NESARC, controlling for Wave 1 demographic characteristics and psychiatric disorders ($p < 0.01$).
- Age

  As shown in Fig. 3.6, the prevalence of current and lifetime alcohol abuse and dependence generally decreases with age. Furthermore, a similar pattern is evident for incident alcohol-use disorders in the

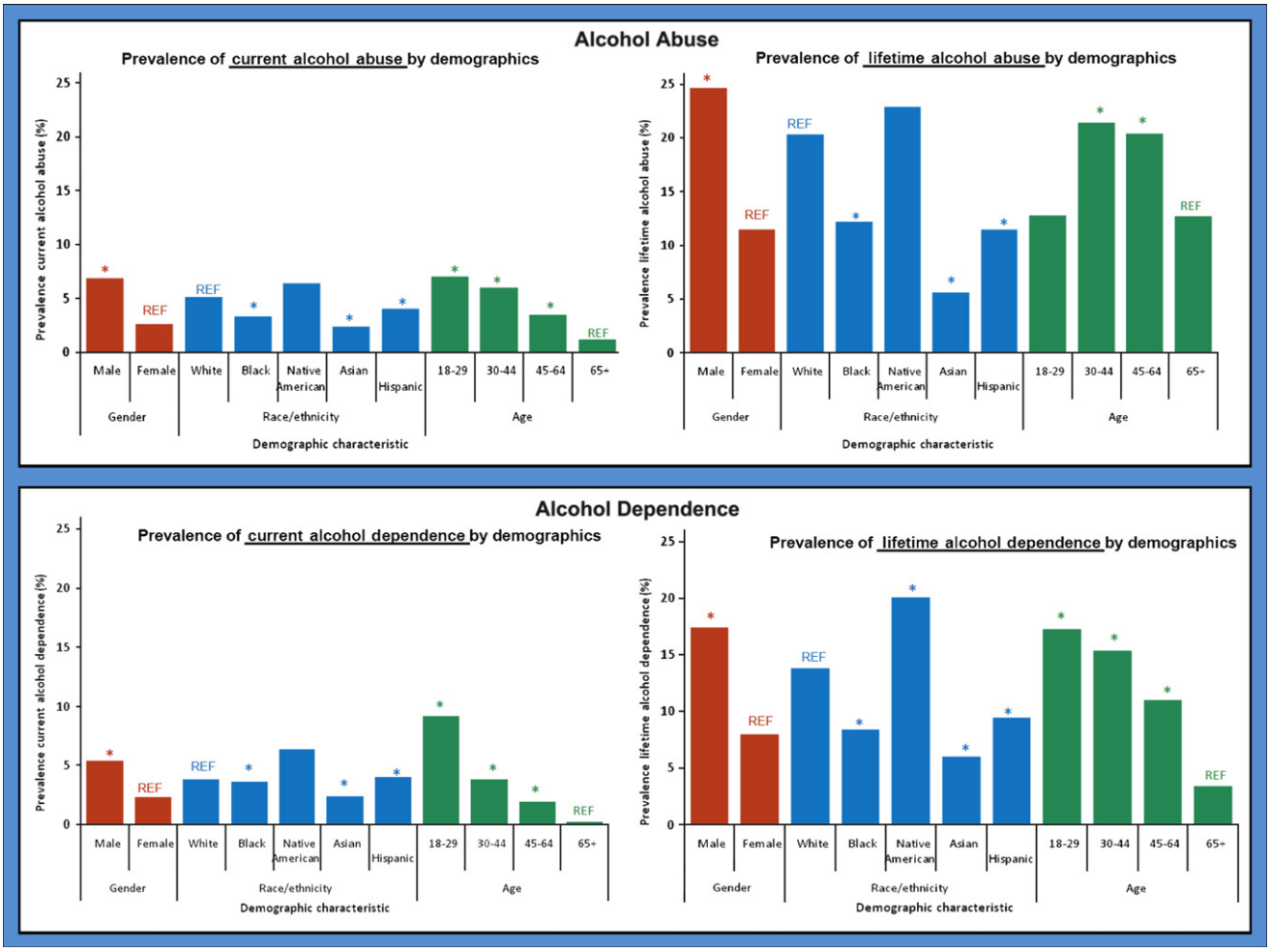


FIGURE 3.6 Sociodemographic correlates of alcohol-use disorders in the NESARC.

NESARC. Specifically, compared to the oldest age group (55 years and older), incident alcohol-use disorders were significantly more prevalent among individuals aged 20–29 (abuse: OR, 11.7; dependence: OR, 8.7) and those aged 30–54 (abuse: OR, 4.3; dependence: OR, 3.5), controlling for Wave 1 demographic characteristics and psychiatric disorders ($p < 0.01$). The finding that younger cohorts were at a higher risk of alcohol dependence and abuse in the NESARC could indicate a true cohort effect or could possibly be the result of under-representation among older cohorts due to differential mortality or poor recall of remote events. Prospective population-based investigation is required to adequately address this issue.

- Environmental factors

Both external and internal environmental factors play a role in alcohol-use disorders. Since alcohol-use disorders can only develop in individuals who have access to alcohol, alcohol availability within an individual's external environment can prevent or facilitate the development of an alcohol disorder. In Western societies, opposing societal forces and social attitudes influences the availability and consumption of alcohol. For example, public health, moral/religious, and governmental organizations try to reduce availability and consumption through public policy and laws, while the alcoholic beverage industry attempts to increase consumption.

Internal environmental factors such as home and family life can also affect the risk of alcohol-use disorders. Childhood environment appears particularly salient, since poor parental monitoring and modeling of heavy alcohol use contribute to the likelihood of adolescent alcohol initiation and binge drinking. Recent epidemiological studies have begun to explore the role of retrospectively reported adverse childhood experiences in the development of adult alcohol disorders.

In the NCS-R, lifetime alcohol abuse was significantly associated with experiencing, as a child,

paternal death, parental divorce or separation, parental psychopathology, repeated sexual molestation, and paternal aggression. There were fewer childhood correlates of lifetime alcohol dependence in this study, with significant associations found only for paternal death and parental psychopathology.

In the NESARC, Thompson and colleagues showed that parental divorce/separation was significantly related to adult lifetime alcohol dependence, after adjusting for parental history of drug, alcohol, depression, and antisocial behavior problems. Pilowsky and colleagues showed that as the risk of adult alcohol dependence increases as a function of the number of childhood traumas experienced (including parental divorce or death, being raised in an institution or by foster parents). Furthermore, Fenton and colleagues showed that childhood sexual abuse, physical abuse, emotional abuse, and physical neglect increased the risk of adult alcohol dependence, independent of experiencing other traumatic childhood events.

Taken together, these findings suggest that it may be important to incorporate an alcohol-use disorder preventative component in harm-reduction interventions conducted among children who experience traumatic events.

- Genetic factors

The important genetic component of alcohol disorders has been well established by findings from population-based epidemiological investigation, familial aggregation studies, adoption studies, and twin studies. At present, the heritability of alcohol dependence is thought to be around 50–60%. Our knowledge of specific genetic causes of alcohol dependence remains limited. However, results from allelic association studies have implicated certain genetic variants including ADH1B*2 (an allele of an alcohol dehydrogenase gene) and Asp40 (an allele of the mu opioid receptor). Understanding the genetic contribution of alcohol-use disorders is critical to the development of effective biological prevention and treatment interventions, and a substantial amount of research in this area is currently underway.

Drugs

- Sex

As with alcohol-use disorders, epidemiological studies show that the prevalence of current or lifetime any drug abuse or dependence is significantly higher among males compared to females (see Fig. 3.7 for NESARC findings). Similarly, studies using NLAES and NESARC data have found that when each drug class is considered specifically, the prevalence of abuse/dependence is higher among males than it is among females. Furthermore, males were significantly more likely than females to develop incident drug dependence (OR, 2.7) (but not abuse) between Waves 1 and 2 of the NESARC, controlling for Wave 1 demographic characteristics and psychiatric disorders ($p < 0.01$).

- Race/ethnicity

Drug use disorders tend to be lowest among Hispanics and Asians and highest among Native Americans; however, racial patterns are less consistent for drug use disorders than for alcohol-use disorders. For example, as shown in Fig. 3.7, in the NESARC compared to Whites Hispanics had significantly lower prevalence of current drug abuse but not of current drug dependence, while Native Americans had significantly higher prevalence of current drug dependence, but not of current abuse. However, lifetime drug abuse and dependence were significantly lower among Hispanics and Asians compared to Whites, whereas lifetime drug dependence (but not abuse) was significantly higher among Native Americans. However, racial differences in nonmedical prescription drug use disorders appear to follow a different trend. The prevalence of nonmedical use disorders was not significantly different for Hispanics, Asians, or Native Americans compared to Whites in the NLAES and NESARC; however, Blacks had significantly lower odds of nonmedical use disorders compared to Whites in both of these studies. Similarly, NSDUH data from 2002 to 2008 reveal that the prevalence of drug abuse/dependence is lowest among Asians and highest among Native Americans/Alaska Natives. For example, in 2008 past year drug abuse/dependence was 4% among Asians, 9% among black, 9% among Whites, and 10% among Hispanics. However, despite these differences in prevalence, there were no significant race differences in the odds of developing incident drug abuse or incident drug dependence between Waves 1 and 2 of the NESARC ($p > 0.05$).

- Age

Epidemiological studies consistently indicate that younger age is associated with an increased prevalence of drug abuse and dependence. For example, Fig. 3.7 reveals that in the NESARC, prevalence of current and lifetime any drug abuse or dependence decreased as age increased and drug specific analyses reveal the same trend for nonmedical prescription drug use disorders and for cannabis use disorders. Likewise, compared to ages 55 years and older, incident drug abuse was higher among ages 20–29 (OR, 50.7) and ages 30–54 (OR, 14.4), and incident drug dependence was higher among ages 20–29 (OR, 8.4), controlling for Wave 1

demographic characteristics and psychiatric disorders ($p < 0.01$).

It is interesting that the two youngest age groups (18–29 and 30–44 years) had almost identical prevalence of lifetime drug dependence (10.1% compared with 10.8%) and very similar prevalence of lifetime drug dependence (4.1% compared with 3.5%) (see Fig. 3.7). Not only are 18–29-year olds generally considered the highest risk group for substance use, but, in addition, epidemiological data indicate that this is the age at which drug use is most likely to begin. For example, in the NCS, the highest risk for initiating marijuana use occurred at around age 18 while the highest risk for cocaine initiation occurred at around age 20. Similarly, in the NCS-R, the median age at onset for any alcohol or drug use disorder was 20 years, within a fairly narrow range of 18–27 years. Thus, the finding that 30–44-year olds do not differ a great deal in terms of drug use from 18–29-year olds may suggest changing social norms and increased acceptance of drug use among older individuals. It is therefore possible that as these cohorts age, population-level increases in drug use among older cohorts may be evident.

- Environmental factors

Unlike alcohol, which can be purchased legally in the United States, it is illegal to obtain drugs for nonmedical purposes. Some neighborhood level factors such as socioeconomic status of residents and prevalence of crime may increase the availability of drugs and therefore increase the risk of drug use among individuals within this environment. Since drug use is a necessary prerequisite for drug abuse and dependence, the risk of these disorders will be increased among individuals living in this environment. Although, to the best of our knowledge, large-scale US-based epidemiological investigation of this phenomenon does not exist, the role of contextual factors in drug use disorders has been increasingly acknowledged since the 1980s.

Many studies have found that low socioeconomic status includes a number of factors that may contribute to a higher likelihood of substance abuse/

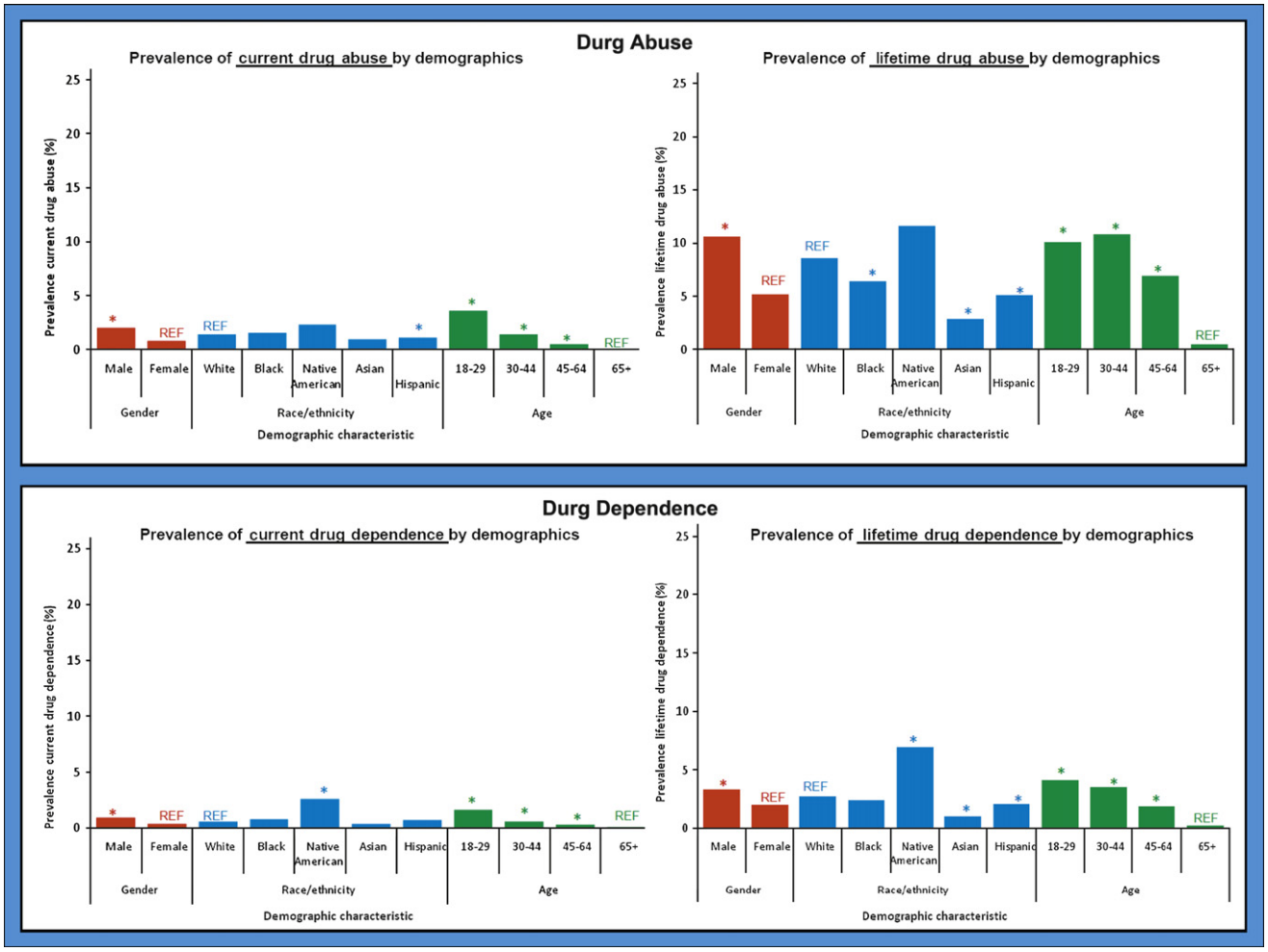


FIGURE 3.7 Sociodemographic correlates of drug use disorders in the NESARC.

dependence. These factors include: (1) residence in impoverished and stressful neighborhoods; these findings support the stress reduction hypothesis, which suggests that drug use disorders are a maladaptive form of coping with hostile environments and negative life events; (2) drug-salient surroundings, which include more drug abusers and drug-abusing social networks; and (3) neighborhood social disorganization, which refers to a lack of cooperation of local residents toward common goals. Neighborhood social disorganization has also been associated with higher rates of local gang activity, which, in turn, is a predictor of drug abuse and dependence. Experiences of discrimination in racial/ethnic minorities have also been associated with increased drug use.

As with alcohol-use disorders, drug abuse and dependence are also thought to be highly influenced by a range of childhood environmental factors. Unfortunately, epidemiological data on the role of childhood adversity in the development of adult drug use disorders are seriously limited. One study using NCS-R data found that lifetime drug abuse was significantly associated with experiencing, as a child, parental divorce or separation, parental psychopathology, isolated and repeated molestation, isolated rape, and maternal aggression, whereas lifetime drug dependence was predicted by parental divorce or separation, parental psychopathology, isolated rape, and maternal aggression. However, as already discussed, the NCS-R did not obtain representative diagnoses of drug dependence.

In light of the gaps in the literature on drug use disorders, it is worth briefly discussing findings on drug use from a large Health Maintenance Organization (HMO) single site study designed to investigate the association between retrospectively reported childhood trauma and adult health outcomes. Using data from this study, Dube and colleagues report that early initiation of illicit drug use (<15 years) was significantly higher among individuals who reported childhood emotional abuse, physical abuse, sexual abuse, emotional neglect, and physical neglect, as a child living with a mentally ill person, witnessing domestic violence, substance abuse, parental separation or divorce, or parental incarceration. Given that NSDUH data indicate that adolescent-onset illicit drug users are at a highly increased risk of developing drug dependence these data are concerning. Future studies to address the relationship between childhood adversity and adult drug use disorders using data including detailed measures of adverse childhood experience and adult drug use disorders a nationally representative sample are required.

- Genetic factors

Genetic epidemiology has established the important influence of genetics in the risk of drug abuse and dependence. Population-based epidemiology studies, clinical studies, and family studies have demonstrated that drug use disorders are familial diseases. Adoption and twin studies indicate that heritability estimates of drug use disorders are somewhat lower than alcohol-use disorders with estimates ranging from around 30 to 60%. However, heritability appears to be higher among males; for example, Kender et al. report that in a sample of 1198 male twins, heritability was 76% for drug abuse, and 69% for dependence.

Association studies involving the dopaminergic and serotinergic systems indicate that drug disorders are positively associated with polymorphisms of dopaminergic receptor genes (DRD1, DRD2, DRD3, and DRD4) and genes associated with serotonin transport (e.g. 5-HTTP). Furthermore, it appears that a combination of several genes may be involved in the development of drug use disorders. However, much remains unknown, and research to identify specific genes and examine interplay between genes and the environment is underway.

CONCLUSION

In contrast to the amount of knowledge on the epidemiology of addiction in the United States available in the 1980s, several large, and nationally representative epidemiological studies have been conducted using standardized diagnostic assessments. These studies provide valuable information on the distribution and determinants of substance use disorders in the United States, and although some variation in prevalence does exist, many of the findings are consistent across studies. It is important to note that interpretation of epidemiological findings should account for possible study limitations such as potential biases stemming from self-report. Nonetheless, epidemiologic studies present the most valid national picture of substance use disorders and findings presented in this chapter and make a significant contribution to our understanding of substance use disorders in the general population.

SEE ALSO

Alcohol Use Disorders, Heroin Addiction, Cocaine Addiction, Marijuana Use and Abuse, Methamphetamine Addiction, Hallucinogens, Ecstasy/MDMA, Inhalants, Ketamine, Anabolic-androgenic Steroid Use and Dependence, Prescription and Over-the-Counter Medications, Tobacco, Gender Differences, The Intergenerational Transference of Addiction, Interpersonal Factors and Addictive

Disorders, Personality and Addiction Processes, Adolescent Substance Use: Symptoms and Course, Symptoms and Course: Alcohol Use Disorder in Adulthood, Symptoms and Course: Older Age and Substance Abuse, Minority Groups and Addictions, Epidemiology of Adolescent and Young Adult Alcohol, Tobacco, and Drug Use and Misuse in the United States

List of Abbreviations

APA the American Psychiatric Association is a scientific and professional organization that represents psychology in the United States. With 150 000 members, APA is the largest association of psychologists worldwide.

AUDADIS the Alcohol-Use Disorder and Associated Disabilities Interview Schedule, a structured diagnostic interview used to measure substance use and mental disorders in the NLAES and NESARC.

CIDI the Composite International Diagnostic Interview, a fully structured interview designed to assess mental disorders and intended for use in epidemiological and cross-cultural studies and for clinical purposes. Used in the NCS.

ECA Epidemiological Catchment Area Study, a multisite study conducted between 1978 and 1983 that obtained DSM-III substance use disorder diagnoses using the Diagnostic Interview Schedule (DIS), with a sample of about 20 000.

MTF Monitoring the Future, an ongoing study that tracks substance use behaviors, attitudes, and values among US adolescents.

NCS the National Comorbidity Survey, a psychiatric epidemiologic survey conducted between 1990 and 1992, with a sample of about 8000.

NCS-R the National Comorbidity Study-Replication, a follow-up study to the NCS conducted between 2001 and 2003, with a sample of about 10 000.

NESARC the National Epidemiologic Survey on Alcohol and Related Conditions, designed to assess alcohol-use disorders and their associated disabilities in the general population. It is a longitudinal survey with Wave 1 conducted in 2001–02 with a sample of 43 093 and Wave 2 in 2004–05 with a sample of 34 653.

NLAES the National Longitudinal Alcohol Epidemiologic Survey, an epidemiologic survey conducted between 1991 and 1992 that assessed alcohol and drug use disorders and other psychiatric disorders, with a sample of 42 862.

NSDUH the National Survey on Drug Use and Health, a large annual national survey of ages 12 and older designed to track trends in substance use.

OR odds ratio.

Glossary

Epidemiology the study of the distribution and determinants of disease in the population.

Incidence the number of new cases of a health outcome that arise in a given population over a certain time period.

Persistence the proportion of cases of a health outcome that continue to experience that health outcome over a certain time period.

Prevalence the number of cases of a health outcome that exists within a given population at a specific point in time.

Reliability consistency and replicability of measurements. Concerns random errors.

Telescoping effect the finding that women typically initiate alcohol use later than men but progress faster to alcohol dependence, and those who become dependent progress faster to treatment. The course of alcohol disorders in women is thus considered to be compressed, or telescoped, compared to men.

validity accuracy of measurements, whether the measure captures the concept that it is intended to measure. Concerns random and systematic errors.

Further Reading

American Psychiatric Association, 1994. Diagnostic and Statistical Manual of Mental Disorders, fourth ed. American Psychiatric Association, Washington, D.C.

Compton, W.M., Thomas, Y.F., Stinson, F.S., Grant, B.F., 2007. Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: results form the National Epidemiologic Survey on Alcohol and Related Conditions. Archives of General Psychiatry 64, 566–576.

Gelernter, J., Kranzler, H.R., 2010. Genetics of drug dependence. Dialogues in Clinical Neuroscience 12, 77–84.

Grant, B.F., Goldstein, R.B., Chou, S.P., et al., 2009. Sociodemographic and psychopathologic predictors of first incidence of DSM-IV substance use, mood and anxiety disorders: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. Molecular Psychiatry 14 (11), 1051–1066.

Grant, B.F., Dawson, D.A., Stinson, F.S., et al., 2003. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. Drug and Alcohol Dependence 71, 7–16.

Hasin, D.S., Stinson, F.S., Ogburn, E., Grant, B.F., 2007. Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Archives of General Psychiatry 64, 830–842.

Johnston, L.D., O'Malley, P.M., Bachman, J.G., Schulenberg, J.E., 2010. Monitoring the Future National Results on Adolescent Drug Use: Overview of Key Findings, 2009 (NIH Publication No. 10-7583). National Institute on Drug Abuse, Bethesda, MD.

Kessler, R.C., Chiu, W.T., Demler, O., Merikangas, K.R., Walters, E.E., 2005. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Archives of General Psychiatry 62, 617–627.

Li, T.K., 2000. Pharmacogenetics of responses to alcohol and genes that influence alcohol drinking. Journal of Studies on Alcohol 6, 5–12.

Regier, D.A., Myers, J.K., Kramer, M., et al., 1984. The NIMH Epidemiologic Catchment Area program. Archives of General Psychiatry 41, 934–941.

Spitzer, R., Endicott, J., Robins, E., 1978. Research diagnostic criteria, rationale and reliability. Archives of General Psychiatry 35, 773–782.

Substance Abuse and Mental Health Services Administration, 2009. Results from the 2008 National Survey on Drug Use and Health: National Findings (Office of Applied Studies, NSDUH Series H-36, HHS Publication No. SMA 09-4434). US Department of Health and Human Services, Rockville, MD.

Susser, E., Schwartz, S., Bromet, E., Morabia, A., 2006. Psychiatric Epidemiology. Oxford University Press, New York, NY.

Relevant Websites

Apparent per capita ethanol consumption for the United States, 1850–2007 (gallons of ethanol, based on population age 15 and older prior to 1970 and on population age 14 and other thereafter). National Institute of Alcohol Abuse and Alcoholism website. Available at: www.niaaa.nih.gov. (accessed 28.09.10.).

CHAPTER

4

An Evolutionary Perspective on Addiction

Elizabeth M. Hill

University of Detroit Mercy, Detroit, MI, USA

OUTLINE

INTRODUCTION

Did the propensity for addiction evolve? If so, what is its possible benefit, and how would it function to increase the likelihood that individuals who have this trait will bear more offspring, thus increasing the frequency of addiction over generations? It seems ludicrous to suggest that addiction, with its high cost to society, is beneficial to any individual or group. Nevertheless, taking an evolutionary perspective may help us understand why and how the human brain evolved to be vulnerable to addiction, and why this negative vulnerability has not been eliminated by evolution. While no one would argue that the state of addiction (i.e. physical and psychological dependence on a drug) has an evolutionary benefit and is therefore an adaptation, the precursors of addictive behavior, such as the propensity to take risks, might actually be advantageous in some settings, and they could be the target of natural selection.

There have been several evolutionary approaches to addiction. The first considers vulnerability to addiction to result from the mismatch between our current and ancestral environments. A second proposes that over time humans developed a tolerance to low levels of psychoactive substances, adaptations that are now overwhelmed by the higher concentrations of drugs available. The third approach explains demographic patterns in problematic substance use with evolutionary forces. Human physiology may have evolved to allow us to ingest plant-based drugs for their beneficial effects, and physiology and brain structure that resulted then became vulnerable to exploitation by potent chemicals, resulting in a disease of modern civilization – substance dependence.

Evolutionary approaches to addiction cannot be tested experimentally, in the traditional sense, because the argument is about past events, and the fossil record is not helpful in this instance. Nevertheless, evolutionary scientists may glean insights from studies of preindustrial societies and also from studies with nonhuman animals. While it is difficult to prove one of these hypotheses over another, a thorough examination will help clarify the needs for future research. To evaluate these approaches to addiction, a review of the basic components of evolutionary theory is useful, as is

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00004-8

a careful examination of the terminology for psychoactive substances. Researchers do not use the term, addiction, in exactly the same way. It is also useful to examine newly emerging areas of evolutionary psychology and Darwinian medicine. Evolutionary scientists use criteria for determining whether to consider a trait to be adaptive; these will be described. After the important concepts of evolutionary theory are introduced, evolutionary approaches to human problems will then be described.

For the purposes of this article, *addiction* and *dependence* will be used interchangeably, and the term *substance consumption* will refer to low levels of use, below the levels that induce intoxication or impairment. Dependence is characterized by tolerance, withdrawal, and compulsive use (inability to abstain despite problematic consequences). Abuse and harmful use involve recurrent use resulting in problems but where dependence has not developed. Binge or risky drinking refers to an episode of intoxication, typically defined as consumption of 4–5 drinks at a sitting. This article will mainly concern alcohol and plant-based alkaloids. Alcohol is usually termed ethanol in a medical context. Plant-based alkaloids include nicotine from tobacco, ephedrine from khat, arecoline from betel nut, and cocaine from coca. The term, drugs of abuse, will also be used to describe the set of drugs with high addiction potential. These are cocaine, alcohol, marijuana, amphetamine, and heroin.

WHAT IS AN EVOLUTIONARY PERSPECTIVE?

An evolutionary approach to a psychological or biological phenomenon entails the analysis of "ultimate" questions, as called by Niko Tinbergen, a pioneering Dutch ethologist. Ultimate questions are about why a behavior exists, i.e. what is the function, and what is the evolutionary history. An evolutionary approach to addiction attempts to address function or evolutionary history or both. Does addiction (or consumption of addictive substances) have any function? What is the evolutionary history of addiction (or consumption of addictive substances) among the ancestors of modern humans? Tinbergen's other questions concerned "proximate" levels, understanding mechanism and development. Proximate questions ask about how, rather than why, a behavior exists. A complete analysis of behavior requires answering these questions, also, but they are beyond the scope of this article.

Evolution is usually defined as a change in gene frequency over generations through the process of natural selection. Charles Darwin and Alfred Russel Wallace explained the process of natural selection, which results when characteristics are passed along from parents to offspring, when organisms show variability, and when there is differential reproduction among offspring. Those who reproduce best are considered the most "fit" for a particular environment ("niche").

The origin of variation is through random mutation of genes. Many mutations are deleterious or even lethal, but some confer advantages. Where a mutation is beneficial, the individual bearing it will be more successful in survival and reproduction. The prevailing view has emphasized selection at the level of the gene, leading to the phrase, "the selfish gene." Recent theories have argued that selection at the level of groups can be seen to explain the human tendency toward cooperation and empathy.

Natural selection includes sexual selection, a process that may have a role in the evolution of some aspects of addiction, to the extent that there are gender differences in patterns of use, abuse, and dependence. Sexual selection pertains to an increase in frequency of sex-linked traits that lead to increased reproductive success. These traits generally are not beneficial for survival chances, but are for reproductive competition, such as antlers on a deer or elaborate plumage on the peacock's tail. Some evolutionary biologists consider sexual selection to be simply a special case of natural selection.

Evidence for past evolution comes from fossils and from comparing characteristics of existing species. Evidence for ongoing evolution comes from longitudinal studies of changes in characteristics over time in response to selection pressures. Evolution can be observed as it happens when selection pressures have changed quickly. For example, Rosemary Grant and Peter Grant conducted a long-term study of Darwin's finches on the Galapagos Islands, during a period when rainfall changed dramatically. The researchers examined the beak depths of seed-eating finches. When rainfall was low, only the birds with the strongest beaks survived and reproduced, because the others could not crack enough seeds for sufficient nutrition. The distribution of beak depths in the population changed within a few generations, illustrating natural selection in action.

Another example of natural selection in action is the rapid development of bacterial resistance to antibiotics over the past 50 years. Infectious diseases were treated very effectively using penicillin and other antibiotics when they were first developed. However, sufficient variation existed in the genetic makeup of bacteria that some survived. Genetic variants or new mutations conferred resistance to penicillin. The resistant organisms survived and reproduced. This evolution has occurred again and again in response to new antibiotics. Staphylococcus infection is a critical problem in hospital

settings, and the latest antibiotics are losing effectiveness against it, as did previous generations of medicines.

A new field called Darwinian medicine attempts to uncover why natural selection has not eliminated specific flaws that lead to disease. The application of an evolutionary approach to the study of psychological and medical disorders most often involves an examination of the history and function of the nervous system's vulnerabilities. Most relevant to addiction, the risk for disease may be increased when previously advantageous mechanisms do not operate properly in the current environment. For example, cardiovascular disease and obesity are prevalent in advanced societies. The perspective of evolutionary or Darwinian medicine examines the ancestral or "natural" environment, where the availability of fat and sugar would never exceed healthy levels, and people had more physical activity. The brain and physiological mechanisms controlling taste, eating motivation, and satiety evolved in an environment totally unlike the current one, in terms of the wide availability and concentrations of these foods. These mechanisms result in a vulnerability to obesity and cardiovascular disease.

COULD ADDICTION EVOLVE?

If a characteristic has evolved, it means that it has become more frequent over generations due to natural selection. Natural selection entails differential reproduction by those who possess traits more favored or "adapted" for a particular environmental setting. For substance use to evolve, it must confer an adaptive advantage for those who have the trait. Thus, to consider substance use to be an adaptation, it must have a clear history of being maintained by natural selection, it must be correlated with fitness during its evolution, and it must have solved an adaptive challenge. If these criteria are not met, then substance use is best considered as a co-opted adaptation (exaptation), a functionless by-product, or a co-opted functionless by-product. An exaptation currently increases fitness; such characteristics are sometimes called preadaptions. Tolerance to low levels of substances could perhaps be an adaptation that increased fitness in the past, but excessive consumption of substances today is clearly a functionless by-product of brain structure and other physiological processes. Although it is clear that substance dependence is not adaptive, some facets of the process of developing an addiction might be, such as tolerance of substances at low levels.

Three main evolutionary approaches consider substance dependence to be a by-product of preexisting adaptations. None of these approaches considers dependence as we see it in contemporary society to be adaptive. These different perspectives treat different aspects of our physiology as creating vulnerabilities.

One evolutionary perspective considers addiction to be a disease of modern civilization. Such a disease results from a mismatch between the conditions prevalent during the evolution of our brain's reward pathway and conditions in the contemporary environment relating to psychoactive substances of abuse. The brain's vulnerability to addiction derives from the fact that neural communication is chemical. Drugs, a kind of artificial and external reward, mimic the natural processes, activating and overwhelming the brain pathway, which becomes dysfunctional. Thus, the brain's reward system is adapted, but addiction to external chemicals is an unfortunate functionless by-product. This perspective is referred to as the "mismatch hypothesis" in this article.

The second major evolutionary approach to addiction looks for features of the addiction process that are adaptive and finds evidence for plant–animal coevolution. In this view, the human capacity to tolerate low levels of certain drugs may have evolved at the same time as plants were evolving defenses against predation by animals. Other primates and mammals, and even insects, have demonstrated a similar response to such substances. Coevolution of traits in two species, resulting in a mutually beneficial relationship, is called mutualism. This approach views substance use at naturally occurring levels as functional, rather than a by-product. In theory, animals have coevolved physiological mechanisms to digest toxic substances for their beneficial effects. This approach is referred to as the "mutualism hypothesis" in this article.

The third evolutionary approach to addiction uses life-history theory, an extension of evolutionary theory. It examines patterns of problematic substance use, positing that an evolutionary approach may help understand individual differences in vulnerability to addiction. Substance misuse and other activities that together fall under the rubric of "risk taking" follow a similar demographic pattern. Evolutionary theory understands the marked differences in risk taking by age and between genders within the framework of sexual selection. Risk taking can be seen as a mating tactic that gives young men an advantage in competing with other young men for partners. Risk taking would be less frequent outside the arena of mating competition. This approach is referred to as the "life-history theory" in this article. These three perspectives are described in detail below.

Addiction as a Disease of Modern Civilization: The Mismatch Hypothesis

This theory posits that drug addiction is a nonadaptive epiphenomenon or by-product of our brain

structure. Psychoactive drugs act on brain pathways that evolved to control emotion and behavior. Humans, along with other mammals, have a "reward" pathway in the brain. Referred to as the mesolimbic system, it is a pathway beginning in the midbrain ventral tegmental area and terminating in a forebrain structure called the nucleus accumbens. The pathway then connects to the amygdala and prefrontal cortex. This pathway is activated when we are engaged in seeking reinforcing events, leading Jaak Panksepp to call this pathway the "seeking" system. It is an appetitive system that is critical for responding to reward and punishment. It supports pleasant and aversive states that are essential for survival, such as hunger. The primary neurotransmitter of this pathway is dopamine, with the key target zones in the nucleus accumbens. Other neurotransmitters, such as opioid peptides, also may be involved in mediating the reward response to positive stimuli.

Recent work in neuroscience has delineated several aspects of reward that are important in effective motivation of behavior. The experience of pleasure ("liking") is separable from "wanting," which is the motivation to seek a reward. The aspect of motivation that seems to be mediated by the mesolimbic dopamine system is "wanting" or "seeking" rather than the experience of reward or pleasure.

Different drugs may stimulate unique profiles of brain changes, but virtually all drugs of abuse activate the dopaminergic mesolimbic system, including heroin, cocaine, alcohol, marijuana, amphetamine, and their synthetic forms. According to Eric Nestler, drugs of abuse increase levels of dopamine in the nucleus accumbens, either directly or indirectly. Direct effects occur with stimulants (e.g. cocaine or amphetamines), which increase the release of dopamine from presynaptic neurons. Alcohol indirectly excites nucleus accumbens neurons by activating gamma amino butyric acid (GABA) neurons, among its other actions. Opiates similarly disinhibit dopamine neurons in the ventral tegmental area by inhibiting GABAergic interneurons. In addition, there are opioid receptors on neurons in the nucleus accumbens, and opiates have a direct effect on these. The mechanism of action of other drugs in this pathway is more complicated.

After chronic exposure to various types of drugs, the mesolimbic dopamine system suffers similar problems. Dopamine reservoirs seem to be reduced, so that natural rewards stimulate less response. In contrast, response to artificial drugs is enhanced or sensitized for an extended time. These effects have been demonstrated in animal studies of cocaine, amphetamines, opiates, alcohol, and nicotine. The body's stress response system is also affected by chronic drug exposure. An additional common effect is reduced activity in the frontal cortex, the region of the brain that exerts higher control over behavioral reactions to activity in the mesolimbic system. With hypofrontality due to chronic drug use, impulsivity is more likely.

These systems evolved through natural selection because they were adaptive; they produce positive emotions in response to stimuli that are important for survival and reproduction (sex, food, positive social contact, and infant attachment). The responses to drugs can be larger than those occurring with natural rewards, so that dopamine may become depleted. Addiction has been considered a brain disease due to this process, in which the evolved mechanism for directing us to seek beneficial rewards is "hijacked."

This perspective has been put forth by Randolph Nesse, a psychiatrist who is one of the founders of Darwinian medicine. Artificial drugs or artificial concentrations of natural drugs will impair the operation of normal motivation and emotion because they falsely signal an imminent benefit to fitness or they falsely block negative emotions that are useful. The brain is vulnerable to false incentives from drugs when their levels exceed those our ancestors encountered in ancestral environments. The potency of today's drugs and their administration through quickly absorbing routes can overwhelm the mesolimbic system. Why has the vulnerability to addiction persisted, rather than being winnowed out through natural selection? This question is easy to answer now, because selection against the mesolimbic system would have highly negative consequences. The mesolimbic reward system has benefited humans and other mammals to an extent that those creatures with such systems do better, survive, and reproduce more, than those who lack such a system. The mesolimbic system enables more effective motivated behavior.

Nesse also asks whether it is wise to block negative emotions such as anxiety by using mood-elevating drugs. The utility of anxiety in the natural environment is well known. However, some negative emotions may no longer be useful, such as panic attacks. Scientific knowledge about the evolutionary function of negative emotions is lacking, and consequently we have no knowledge of the possible wisdom of medicating or self-medicating such feelings.

According to Eric Nestler, particular drugs can be discriminated, primarily through their effects beyond the dopaminergic mesolimbic system. Effects on other parts of the nervous system and on the cardiovascular system differ among various drugs. In addition, there are differences in the effects of chronic use on specific receptor systems that are directly stimulated or antagonized by various drugs, such as the effects of opiates on opioid receptors and nicotine on nicotinic cholinergic receptors. The fact that drug profiles are not identical

has led to criticism of a unitary reward model for drugs of abuse.

Plant–Herbivore Coevolution: The Mutualism Hypothesis

The mismatch theory described above contends that drugs available in current environments are novel and overwhelm evolved mechanisms. In contrast, another evolutionary approach argues that there may be a long history of human substance use, that our exposure to psychoactive chemicals is not a recent development.

The use of psychoactive substances by current hunter-gatherer groups has been documented. Written records and archeological evidence trace the cultural use of drugs back about 10 000 years. The most common drugs in the world today are caffeine, ethanol, nicotine, and betel nut. Alcohol is the most commonly recorded substance used in earlier civilizations, from ancient Egyptians and Greeks to ancient China and India. The use of betel nut can be dated to over 10 000 years ago, and use of coca plants to over 5000 years ago; both of these plants have stimulant effects. Fermentation was invented about 9000 years ago in prehistoric China, according to some sources, but the fermentation of alcohol from grain, honey, or fruit can also be traced to ancient Egypt, Greece, and India.

Early human ancestors, other primates, indeed other mammals, would have access to plants as food. Fruit-eating, frugivory, is thought to expose animals to low levels of ethanol, and plant-eating, herbivory, is thought to expose them to low levels of various chemicals. Fruit-eating relates to the history of alcohol use, and eating chemicals produced in leaves and seeds relates to the use of a number of other drugs.

Fruit and Ethanol

Robert Dudley, a biologist, proposed that an attraction to ethanol might have a function. In the ancestral lineage leading to humans, fruit-eating (frugivorous) monkeys and apes (anthropoids) predominate. Our closest ancestors are primarily frugivorous, the hominoids (gorilla, chimpanzees, orangutans, and gibbons). Subsisting on fruit entails some ingestion of fermenting fruit, and the anthropoid lineage may have been exposed to low levels of ethanol through eating fermented fruit. Yeasts ferment plant sugars to yield ethanol. Microbes compete for access to plant sugars, and they are common within plant reproductive organs (of which fruit is part). It has been suggested that the adaptive function of fermentation by yeast was the killing of bacterial competitors for fruit sugars. Bacterial growth is inhibited by ethanol at levels that can be tolerated by yeast. Ethanol within ripe and fermenting fruit is a natural and common occurrence (various reports measured 0.5–1.6% ethanol in pulp of ripe and 0.6–2.6% in overripe fallen fruit). Decomposition of fruit by fermentation is not desirable to the plant's reproduction, however, if ingestion and dispersal of seeds is deterred by decomposition.

Dudley theorized that the ubiquitous presence of ethanol in ripe fruit could enable its use as a cue for frugivores. Ethanol could serve various roles as a cue. For example, ethanol odors could help animals locate ripe fruit, and ethanol could then take on a role as an appetitive stimulus to rapid consumption of a transient nutritious resource. At present there is no direct evidence for this hypothesis, and the potential uses of ethanol as a cue would pertain for any fruit-eating animal.

A physiological response to ethanol could have been retained in humans as we evolved a more diverse omnivorous diet over the past 2 million years. Humans probably consumed a lot of fruit until the advent of agriculture about 10 000 years ago. These previously advantageous behaviors could represent a vulnerability to negative effects when humans are exposed to the much more concentrated and more widely available levels of ethanol in the present day contemporary environment. This argument is similar to Nesse's mismatch theory described above. Exposure to higher concentrations of ethanol than are naturally available (0.5–1.6% ethanol in pulp of ripe and 0.6–2.6% in overripe fallen fruit) may cause harm, whereas metabolic capacity to handle the natural levels would be favored by evolution.

A recent report supports the potential role of the primate diet in vulnerability to addiction. Species of small mammals in Malaysia were documented to specialize on eating fermented floral nectar, which contains ethanol. The bertam palm has flower buds that are home to fermenting yeast. The alcohol concentration in nectar samples averaged about 0.5% and was as high as 3.8%. The small mammals pollinate the palm through their nectar ingestion, carrying pollen grains on their hair from flower to flower. This situation is an example of mutualism, where coevolution has undoubtedly occurred.

The mammals seen to drink fermented nectar included pentailed treeshrews, a primitive mammal that is similar to extinct species found over 55 million years ago. It is in the lineage ancestral to primates. Other species considered likely to consume substantial amounts of alcohol through nectar included the common treeshrew, plantain squirrel, and slow loris (a larger prosimian primate). Hair analysis of ethyl glucuronide, an ethanol metabolite, showed evidence of chronic ethanol consumption in the plantain squirrel, common treeshrew, and pentailed treeshrew. Loris hair was not sampled in this study. The actual blood alcohol

concentration reached when drinking palm nectar is unknown, however. Bertam palm nectar seemed to be the preferred food source of the pentailed treeshrews and slow lorises.

Despite consuming levels of ethanol that would typically be dangerous, these mammals showed no obvious signs of intoxication to observers. They seem to have evolved sufficient physiological tolerance to allow their exploitation of palm flower nectar as a food source. They can be considered adapted to a unique ecological niche. Thus, such adaptations are within the realm of possibility for other animals. The physiological basis for consumption of nectar might serve as a preadaptation or exadaptation as defined above. Given the ancient presence of these mammals, the relevant physiological accommodations to ethanol could be widespread in the mammalian lineage. However, this level of specialization on an alcoholic food source is rare.

Plant Allelochemicals

Robert Sullivan and Edward Hagen, two anthropologists, provide historical and archaeological evidence that substance use was common among hunter-gatherer, indigenous groups before contact with Westerners. The plants commonly utilized included tobacco (nicotine), khat (ephedrine), betel nut (arecoline), pituri (nicotine), and coca (cocaine). These plants were eaten as food. Plants have evolved defenses against consumption, in the form of alleochemicals that mimic neurotransmitters and bind to receptors in the nervous system of mammals that consume plants (herbivores). Plant neurotoxins disturb the nervous system, structurally or functionally. Herbivores, in turn, have evolved defenses against these toxins that enable continued consumption. For example, humans and other primates eat clay and charcoal to bind the dietary toxins. It is a paradox that these types of drugs from plants are not pleasant on ingestion; rather they are aversive, yet people and other animals consume them, nevertheless. Sullivan and Hagan proposed that mammals evolved adaptations to enable consumption of allelochemicals for their beneficial effects.

Plants and herbivores have undergone a long evolutionary conflict. An example is nicotine, from a domesticated North American tobacco plant. At least 20 different herbivore species of insects and mammals eat this plant. The presence of nicotine in leaves has been shown experimentally to deter ingestion. Countermeasures evolved by herbivores include metabolic processes and enzymes that counteract or detoxify the toxins (or sequester them) and mutualistic relationships with microbes that can allow the extraction of nutrients.

The presence of metabolic physiology implies that humans had regular exposure to plant toxins. The cytochrome P450 (CYP) proteins apparently serve as detoxification enzymes. These are very common throughout the animal kingdom. Their central function is in fatty acid metabolism and synthesis of cholesterol and steroids. Other forms of CYP metabolize drugs (CYP 2 and 3). Some variants of CYP genes may be more common in geographic areas where exposure to certain plants is more common. There may have been population-specific directional selection for the polymorphism CYP2D6 in North East Africa, Turkey, and Saudi Arabia, where khat and the opium poppy are ingested. Such variation in gene frequency is suggestive of fairly recent exposure to different ecologies, while the fact that much of CYP function is conserved suggests older and continued function in more universal metabolic pathways.

Sullivan and Hagen admit that the phenomenon of addiction to modern euphoric drugs might be due to an evolutionary mismatch but they contend that the use of these drugs is far less frequent than another category they call mundane drugs (tobacco, cannabis, and betel nut). Their interest is to create a general theory of drug use based on mundane drugs, rather than the rarer euphoric drug use. Tobacco and cannabis are used by far more people than are heroin, cocaine, and amphetamines.

Plant toxins may have some beneficial effects when ingested by mammals. The toxins present in plants may be useful in herbivores as protection against parasites. Studies of the tobacco hornworm revealed that parasitism by a wasp was less likely for larvae that had ingested nicotine. The hornworm had developed a counter measure to use the plant toxin. Toxic plants may be useful against intestinal parasites. Farmers use tobacco and arecoline from betel nut as an antiworm (helminth) treatment for cattle and sheep. A human example is our use of spices; it has been proposed that spices deter bacterial infection.

Our contemporary society has manufactured better medications for helminth infections than tobacco, but it is plausible that more natural populations might benefit from physiological tolerance to the tobacco plant to exploit its medicinal properties against parasites. Another potential beneficial effect of drugs derived from plant toxins is analgesia. Marijuana was used for medicinal purposes in ancient China, and opium from poppy flowers was used in eastern cultures (Islamic cultures, India, China, the Middle East) for its analgesic effects.

Various hallucinogens and stimulants were used in religious rituals by hunter-gatherer peoples. The Incas chewed coca leaves containing cocaine during religious ceremonies. Hallucinogens have been used for centuries in religious ceremonies, such as psilocybin from mushrooms, which was used by the Aztecs in Mexico and Central America. Psychoactive mushrooms may have been used in ancient India. These substances are thought

to have been used to augment a trance state and induce mystical experiences during religious rituals. Further research is needed to evaluate whether any of these beneficial effects might have a role in the evolution of metabolic processes to utilize these plants, to support the hypothesis of coevolution. Unfortunately, the practices of earlier societies are exceedingly difficult to study.

Demographic and Geographic Patterns in Addiction: Life-History Theory

Extensive research has focused on assessing individual differences in risk factors for drug and alcohol addiction. Much clinical research has this focus: vulnerability factors in the biological and psychosocial realms. These literatures are beyond the scope of this article, which concerns the characteristics of all humans as a species, rather than differences among individuals. There are, however, two areas of research where an evolutionary perspective is relevant.

First is the area of geographic variation in alcohol metabolism. There are geographic patterns in metabolic enzymes that covary with a history of exposure to alcohol. Response to alcohol is affected by variations in the liver enzymes alcohol dehydrogenase (ADH) and acetaldehyde dehydrogenase (ALDH). Ethanol is converted to acetaldehyde by ADH, then ALDH catalyzes acetaldehyde to acetate. The characteristics of metabolism can differ depending on which of the several variants of these enzymes are present. With slower-acting ADH enzymes, less of the aversive acetaldehyde accumulates, and where ALDH is of a slow form, toxic acetaldehyde accumulates. A faster form of ALDH can clear it more quickly. An accumulation of acetaldehyde is associated with an aversive reaction including nausea and flushing. With slower-acting ADH enzymes, the experience of drinking is more pleasant (euphoric, anxiolytic). Genetic studies have shown that risk for alcohol dependence is associated with variants of ADH and ALDH. In one such study of Japanese and Taiwanese alcoholic patients, the frequencies of faster acting forms were lower for ADH and higher for ALDH, compared with nonalcoholic comparison groups.

The geographic distribution of alcohol-metabolizing enzymes can be understood using an evolutionary approach such as the one used to explain the spread of lactase persistence. Natural selection pressures can maintain genetic variation where there are environmental differences. Mammalian milk contains the sugar lactose. Infants digest lactose until weaning, but then stop synthesizing the necessary enzyme. After that stage, gastric distress will result when fresh milk is consumed. Exceptions to this norm are found in northern European, North African, and Arabian populations, who have lactase persistence as adults. These populations are also ones who practice dairying but not farming. The advantages of an ability to digest fresh milk are obvious in these ecological settings. Interestingly, lactase persistence evolved in Europe and Africa by independent mutations in different genes for the metabolic enzyme.

A similar analysis could explain patterns in the geographic distribution of genetic variants in alcohol-metabolizing enzymes. A team of Chinese scientists has examined 38 populations across East Asia, comparing the prevalence of one specific ADH polymorphism to geographic variation in diet. In particular, they looked at the spread of rice cultivation, which emerged about 10 000 years ago. There is a gradual shift from east to west, as shown by cultural relics. Southern and south-eastern China grew rice earliest (10 000 years ago), followed by central China (3000–6000 years ago), with Korea and Japan domesticating rice within the past 3000 years. Fermentation of rice is thought to have followed rapidly afterward, beginning about 9000 years ago in southern China. The initial dietary use is thought to be for food preservation. Variation in enzymes related to alcohol metabolism would enable some individuals to utilize fermented foods more readily.

The 38 populations show different frequencies of the particular ADH variant, ranging from 98.5% to a low of 1.7%. This ADH variant metabolizes ethanol faster than the ancestral variant. The statistical correlation of the prevalence of the derived, evolutionarily newer, allele with the geographic order of rice-culture origins was a strong one ($r = .769$, $p < .01$). This pattern cannot be explained by random genetic drift. Rather, the results are consistent with natural selection for tolerance to ethanol that resulted in differences among ethnic groups.

The second area of research on individual differences where an evolutionary perspective is relevant is focused on patterns of drug use that may result from selection pressures. Gender and age patterns in substance misuse follow a similar pattern as many risky behaviors, such as reckless driving and physical fighting. Excessive drinking or illegal drug use may be an outgrowth of risk taking in general. Risky drinking is usually defined as consuming more than five drinks at a sitting for men and four drinks for women. This level of consumption is typically intoxicating to most drinkers. A number of research studies have reported gender and age differences in rates of risky drinking, alcohol abuse, and alcohol dependence.

The highest rates of alcohol problems are shown by young men, measured using lifetime alcohol abuse or dependence over the lifetime, past year, or past 2 weeks.

Rates for men are 3–4 times as high as those for women, throughout the lifespan. Surveys of college students report higher rates of risky drinking by men than women. Similarly, men report higher rates of driving while intoxicated.

With regard to age, the onset of alcohol abuse is low during childhood and early adolescence, peaks during the ages 15–29, and declines thereafter. The rate of onset becomes negligible after age 35. Diagnoses of alcohol abuse and dependence reach the highest rates between ages 18 and 29 for both men and women, decreasing at older ages. Epidemiological studies of these diagnoses have found that prevalence decreases steadily to less than 1% for those over age 80. Risky drinking (five or more drinks/occasion) and weekly intoxication rates were highest for men aged 18–29, steadily decreasing at older ages. Women showed lower rates of these consumption patterns overall, with higher frequencies during ages 18–49, decreasing with age.

Another demographic pattern relates to marriage and parenthood. Risky drinking and dependence become less frequent after people marry and start families. Rates of alcohol abuse/alcohol dependence are consistently lowest for married people and highest for those who are unmarried (never married or separated/divorced), as reported by national surveys. College students follow the same trend, with married students being half as likely to consume alcohol at risky levels. Alcohol-impaired driving shows the same difference. Longitudinal studies that follow individuals have determined that declines in alcohol use occur after marriage and parenthood, rather than preceding these events.

Elizabeth Hill and Krista Chow used life-history theory to understand these demographic patterns. Life-history theory examines biological characteristics of the life cycle (e.g. age at maturation, average number of offspring). According to this theory, these hallmarks are subject to natural selection and reflect adaptation to ecological circumstances such as predation rate. Biological effort is finite, and evolutionary pressures determine the best allocation mix to the tasks of survival, current reproduction, and future reproduction. Gender and age-related differences in risk taking can be understood from this perspective.

The gender that is more specialized in parenting will necessarily devote less effort to mate attraction and retention. If one gender specializes, the other gender is freer to engage in mating with multiple partners. If the females of a species devote more effort to parenting, then males can compete among themselves for multiple mates. This situation leads to higher variability in reproductive success for males than for females. Some males are not able to mate at all, while others monopolize more than one mate. In this situation, female outcomes would be less variable and would not depend on competing with other females for more sex partners. Human males in many societies are faced with this situation, where some are excluded from reproduction, and taking great risks is required for success. When status and resources are at stake, men are more likely than women to engage in physical conflict. This phenomenon has been called the "young male syndrome" by the evolutionary psychologists Margo Wilson and Martin Daly. During the life stage of competing for mates, the potential gain from risk taking is great, but after one reaches a stage of successful mating and is actively engaged in parenting, risk tolerance is expected to decline. Human decisions about status, reproduction, and risk taking are not necessarily conscious; the ultimate consequences are of interest from an evolutionary perspective.

The idea that impulsive risk taking could be functional, i.e. adaptive, is more clearly seen when examining other animals. Extensive studies of rhesus macaques over many generations have shown that alcohol consumption is correlated with a propensity toward impulsivity. The underlying trait of impulsivity may be beneficial in certain contexts, according to the primatologists Melissa Gerald and Dee Higley. The fact that impulsivity is associated with increased alcohol consumption is thus an unfortunate consequence of a trait that has been beneficial in evolutionary terms. The benefit occurs because the genes that contribute to increased mortality through impulsive risk taking may also increase the probability of reproducing. To reproduce, male rhesus monkeys migrate to other groups. Risk-averse, less aggressive males may not be able to defend themselves during migration, or they may fail to leave the natal group. Genes with pleiotropic effects, beneficial at a young age but deleterious later, are difficult for natural selection to eliminate. With increasing age, the power of natural selection declines, relative to its action before reproduction. If the same genes increase the probability of early reproduction but decrease the likelihood of survival later, natural selection could favor these genes under some conditions. Impulsivity could thus create vulnerability for excessive consumption of alcohol and other substances, which has deleterious effects in the current environment of humans.

CONCLUSION

Addiction is not adaptive, but vulnerability to become addicted resulted from natural selection for other traits. An evolutionary perspective can help understand this vulnerability. Three main evolutionary

approaches were reviewed, the mismatch hypothesis, the mutualism hypothesis, and life-history theory. Vulnerability to addiction may result from the mismatch between our current and ancestral environments. Alternately, over time humans may have developed a tolerance to low levels of psychoactive substances present in plants. Lastly, life-history theory aids our understanding of demographic patterns in problematic substance use. Each approach has weaknesses, but each also suggests a number of areas for future research.

The weakness of the mismatch hypothesis is the incomplete state of knowledge about common actions of drugs on brain reward systems. Which is more important, the unique drug profiles or the common actions of drugs? Regardless of the eventual resolution of this question, it is apparent that treatments or medications for addiction should not interfere with the mesolimbic reward pathway, to avoid damaging our responses to natural rewards.

The weakness of the mutualism hypothesis is that coevolution requires that mammals benefit from ingesting toxins or rotting fruit. The nutrition derived from plants eaten is a clear benefit but are there benefits from the ethanol or toxins themselves? Perhaps plant toxins help control parasites, but much more research is needed in this area.

The weakness of the life-history theory approach is its lack of practical or clinical applicability. While this perspective can synthesize a number of patterns in use and abuse of alcohol and other drugs, it does not easily generate ideas for improving treatment.

In conclusion, humans and other animals consume plants for nutrition, and our mammalian ancestors underwent a coevolutionary arms race with plants to enable digestion and metabolism of the antiherbivore chemicals they produce and the other microbes and organisms that compete with us for fruit pulp. Humans and other animals became adapted to tolerate some alcohol and some neurotoxins that must be ingested to survive on plants. These adaptations left us vulnerable to addiction when high dosages became widely available. Chemicals that affect the brain's natural reward system are the most addictive. The mesolimbic system is ancient, but the potent drugs available now are novel compared with natural rewards or naturally available concentrations. Our ability to distill and concentrate doses is a modern phenomenon. It is also new, in terms of mammalian evolution, to administer drugs through routes with faster absorption than oral ingestion. The motivational and emotional systems involved are essential and adaptive in most environments, so selection cannot act to eliminate these pathways. Unfortunately, humans will probably remain vulnerable to addiction.

SEE ALSO

Medical Toxicology of Drugs of Abuse, Developmental Risk Taking and the Natural History of Alcohol and Drug Use among Youth

Glossary

Alkaloids nitrogen-containing organic compounds produced by plants, such as nicotine, morphine, caffeine, and atropine.
Analgesia decrease or relief of pain sensation without loss of consciousness.
Cytochrome P450 a superfamily of liver enzymes that are involved during breakdown of drugs and toxins.
Dehydrogenase an enzyme that catalyzes the removal of hydrogen atoms in physiological reactions.
Helminth a parasitic worm, including flukes, tapeworms, and roundworms.
Herbivore animal whose diet consists entirely of plants and plant products.
Satiety a state of feeling satisfied.

Further Reading

Buss, D.M., Haselton, M.G., Shackelford, T.K., et al., 1998. Adaptations, exaptations, and spandrels. American Psychologist 53, 533–548.
Dudley, R., 2002. Fermenting fruit and the historical ecology of ethanol ingestion: is alcoholism in modern humans an evolutionary hangover? Addiction 97, 381–388.
Gerald, M.S., Higley, J.D., 2002. Evolutionary underpinnings of excessive alcohol consumption. Addiction 97, 415–425.
Grant, P.R., Grant, B.R., 2008. How and Why Species Multiply. Princeton University Press, Princeton, NJ.
Hagen, E.H., Sullivan, R.J., Schmidt, R., et al., 2009. Ecology and neurobiology of toxin avoidance and the paradox of drug reward. Neuroscience 160, 69–84.
Hill, E., Chow, K., 2002. Life-history theory and risky drinking. Addiction 97, 401–413.
Lende, D.H., Smith, E.O., 2002. Evolution meets biopsychosociality: an analysis of addictive behavior. Addiction 97, 447–458.
Levey, D.J., 2004. The evolutionary ecology of ethanol production and alcoholism. Integrative and Comparative Biology 44, 284–289.
Low, B., 2000. Why Sex Matters: A Darwinian Look at Human Behavior. Princeton University Press, Princeton, NJ.
Nesse, R.M., Berridge, K.C., 1997. Psychoactive drug use in evolutionary perspective. Science 278, 63–66.
Nesse, R.M., Williams, G.C., 1994. Why We Get Sick: The New Science of Darwinian Medicine. Times Books, New York.
Nestler, E.J., 2005. Is there a common molecular pathway for addiction? Nature Neuroscience 8, 1445–1449.
Newlin, D.B., 2002. The self-perceived survival ability and reproductive fitness (SPFit) theory of substance use disorders. Addiction 97, 427–445.
Panksepp, J., Knutson, B., Burgdorf, J., 2002. The role of brain emotional systems in addictions: a neuro-evolutionary perspective and new 'self-report' animal model. Addiction 97, 459–469.
Peng, Y., Shi, H., Qi, X., Xiao, C., Zhong, H., Ma, R.Z., Su, B., 2010. The ADH1B Arg47His polymorphism in East Asian populations and expansion of rice domestication in history. BMC Evolutionary Biology 10, 15.
Reznik, D.N., 2010. The Origin Then and Now: An Interpretive Guide to the Origin of Species. Princeton University Press, Princeton, NJ.

Sullivan, R.J., Hagen, E.H., 2002. Psychotropic substance seeking: evolutionary pathology or adaptation? Addiction 97, 389–400.
Tinbergen, 1963. On aims and methods of ethology. Zeitschrift Fuer Tierpsychologie 20, 410–433.
Wiens, F., Zitzmann, A., Lachance, M.-A., et al., 2008. Chronic intake of fermented floral nectar by wild tree shrews. Proceedings of the National Academy of Sciences 105, 10426–10431.
Wilson, M., Daly, M., 1985. Competitiveness, risk taking, and violence: The young male syndrome. Ethology and Sociobiology 6, 59–73.

Relevant Websites

http://evanthsoc.org – Evolutionary Anthropology Society.
http://evmedreview.com – Evolution and Medicine Review.
www.hbes.com – Human Behavior and Evolution Society.
www.becominghuman.org – Institute of Human Origins.
http://evolution.berkeley.edu/ – Understanding Evolution.

CHAPTER

5

International Perspectives on Addiction

Jo Thakker

University of Waikato, Hamilton, New Zealand

OUTLINE

DRUG USE ACROSS TIME AND PLACE

Some researchers have speculated that drug taking may have been a normal part of life at the time that our species was emerging, about 120 000 years ago. However, in terms of concrete evidence, we are limited to more recent history. For example, excavations in modern-day Iraq indicate that both wine and beer were routinely given to Mesopotamian soldiers 5500 years ago and recipes for beer have been found in 5000-year-old Mesopotamian clay tablets. It is well known that wine has long played a prominent role in Mediterranean cultures and its use in these cultures dates back at least 2000 years. Also, Egyptian papyri dating to 1700 BCE contain many references to alcoholic drinks, including regulations regarding their consumption. Further, regulations around the consumption of alcohol were at times seen in both ancient Rome and ancient Greece.

Research suggests that other psychoactive substances were also used in ancient societies. For instance, evidence from Neolithic sites in modern-day Switzerland indicates that opium poppies were grown in this area more than 6000 years ago. There is also evidence that ancient Egyptians used opium poppies as analgesics and sedatives to manage temperamental infants. Similarly, it appears that in both ancient Greece and ancient Rome opium was widely used for medicinal purposes. For example, opium is referred to in Homer's "Odyssey," wherein it is described as a substance that can "quiet all pains and quarrels" (Scarborough, 1995). Also, opium is frequently referred to in the prominent "Materia Medica" (written by Dioscorides in about 70 CE) which suggests that it could be used for a range of ailments, including pain, digestive problems, and insomnia.

The use of cannabis also appears to date back many years. There is evidence that it was used in China at least 6000 years ago and in India about 4000 years ago. Although, it is unclear whether its early use in China was as a drug or as a source of fiber. But, in both locations, it was clearly used for medicinal purposes at some points in history. Further, it appears that it was also used in religious ceremonies in India. Many years later the Greek historian Herodotus described a purification ritual (in the fifth century BCE) conducted by Scythians who were nomadic tribes that inhabited an area that incorporates modern Ukraine, Russia, and Asia. Apparently, the Scythians threw hemp seeds onto hot stones inside a tent so that they could inhale the intoxicating smoke.

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00005-X

It can be concluded then that drug use has been present throughout much of human history; however, it is unclear whether addiction was present or common in early human societies. Given the addictive properties of some of the substances that were used historically, such as opium and alcohol, it is possible that drug dependence would have occurred if the substance was available in sufficient quantities. It is probably impossible to draw any definitive conclusions about the supply of such substances, although, clues may be seen in the way that communities lived in ancient times. For instance, a regular supply of opium was probably not available prior to the advent of agriculture about 10,000 years ago. Also, in many ancient human societies, life was tough and individuals who typically lived within small social groups were expected to contribute to the group. It is hard to imagine that drunks would have been tolerated in these sorts of social contexts as their level of contribution to the group would likely have been severely limited. Thus, it may be the case that individuals would not have allowed themselves to succumb to addiction because of the high personal and social costs that would have been associated with the development of substance dependence.

Today drug use and addiction are seen in most societies, although the substances differ across different cultural settings. Kava (a beverage made from the roots of the kava plant) is commonly used in the Pacific region, whereas betel nut (the seed of the areca palm) is commonly used in India and Southeast Asia. While substances such as these are used in specific regions, other substances, such as tobacco and alcohol are used around the world by almost all cultural groups. One particularly important point to note is that in many countries the drug of choice has changed over time. This is particularly true in countries that have been colonized as often the colonizing country brings with it particular drug use practices that greatly influence the drug use habits of the indigenous people. Australia provides an excellent example of this process.

Prior to European colonization, aboriginal Australians used a number of substances, including, alcohol, native tobacco, and pituri (a native shrub). Interestingly, although alcohol was made from a variety of native plants, its use was apparently uncommon due to difficulties with supply and transportation. It seems that the most commonly used substance prior to the arrival of Europeans was pituri which is a hallucinogen. It was used in a variety of ways, including as an anesthetic, an amnesic, and as a means of inducing altered states of consciousness during various ceremonies. It has been widely argued that one of the reasons that substance use, such as the Aboriginal use of pituri, did not become substance abuse, was that its use was rigidly controlled by social norms and traditions. In other words, people accepted the predominant social view pertaining to the appropriate use of the drug.

Today, indigenous Australians are overrepresented in drug abuse statistics. For example, there are widespread problems among these populations with the use of alcohol and inhalants, both of which are highly addictive. Furthermore, unfortunately, in regard to inhalants, children and adolescents are especially vulnerable. Evidence indicates that these substance abuse problems are closely associated with socioeconomic issues, such as poverty and poor housing, and also with social isolation. Furthermore, those with abuse and dependence problems are also more likely to end up in the criminal justice system.

It has been suggested that one of the reasons that novel substances become a problem in indigenous populations is because they lack a social structure that might control the use of the substance. For instance, in many Western societies, alcohol is used in specific settings, such as special occasions or after regular weekly events (such as a Friday after work). Another example of this is the fact that most people would not consider drinking before the evening, except perhaps on special occasions such as Christmas. In this way, alcohol use has a certain place within an individual's regular schedule and this assists in limiting and controlling its use. In contrast, populations that have been introduced to the substance more recently have no social structure for controlling its use and thus they are vulnerable to developing substance use problems.

A similar picture is seen in New Zealand and Canada; both of which have a colonial history. In both regions, indigenous populations are overrepresented in drug use and addiction statistics. Furthermore, in both nations, indigenous populations are more likely to live in impoverished environments, to be unemployed, and to have poor health outcomes. Thus, substance use problems are closely associated with a range of other life challenges. Subsequently, it is has been widely theorized that addictive behaviors arise, at least in part, because of general life difficulties. Using a substance provides a means of psychological escape from the challenges of daily life. Also, in some instances, substance use and even addiction is normalized, such that it is difficult for a young person to choose a different path.

EVOLUTIONARY AND BIOLOGICAL EXPLANATIONS

Evolutionary explanations are important in psychology because they purport to have relevance for all human groups; in other words, they attempt to explain general human behavior. Evolutionary

psychologists propose that the mind of modern humans is comprised of various psychological mechanisms that evolved in ancient times due to the fact that they promoted survival and therefore enhanced reproductive fitness. Evolutionary psychologists argue that behavior in modern humans may be understood by analyzing these mechanisms because the mechanisms have changed little since modern humans evolved. Thus, even though the environment has changed markedly, the human brain and its mechanisms are largely unchanged. It is important to note that evolutionary psychologists offer distal explanations that refer to adaptation and phylogeny. As such, they are generally complimentary to (and not in competition with) other more proximal explanations.

In light of the ubiquity of drug taking, it is reasonable to ask whether such behavior is due to the presence of a particular adaptive psychological mechanism. In considering this question, it is important to acknowledge the distinction between adaptations and by-products of adaptations. While the former are generally a direct result of the process of evolution, the latter are simply by-products of that process. For example, the kidneys' key function is to filter waste from the blood and they are adapted to this specific purpose. However, they also have other properties such as their color, which serve no real purpose but which have nonetheless arisen through the same natural selection process.

There are a number of ways in which drug use may have boosted reproductive fitness in ancestral environments. Stimulants such as tobacco and coffee have the effects of decreasing the need for sleep and suppressing one's appetite. Analgesics such as opium can be used to relieve pain and discomfort. Thus, it is possible that early humans' chances of survival would have been increased if they used substances for these purposes. However, if this was the case, it is unclear exactly what would have been selected for. For instance, it may simply have been the case that this sort of use would have been more likely to have been utilized by smart individuals, and thus, all that would be passed on in this case was smartness.

With regard to ethyl alcohol, a clear case has been made for the role of a more specific psychological mechanism. It has been suggested by some researchers that we have evolved a preference for ripe fruit because eating such fruit provided an important source of energy that was able to be metabolized quickly. Such fruit typically contains significant quantities of ethanol. This is reflected in the fact that in many cultures throughout human history, alcohol has been viewed as an important source of food. Thus, some researchers have suggested that human beings have evolved a tendency to view ethanol as nutritionally rewarding. Interestingly, in backing up this claim some researchers have pointed to the myriad of examples of this sort of association in the animal kingdom. For example, monkeys, elephants, and butterflies have all been observed to consume overripe fruit to the point of intoxication.

One important issue when considering evolutionary explanations is the mismatch between ancestral and modern environments. What may have been adaptive in a primitive preagricultural setting may not be adaptive today. For instance, a 100 000 years ago, a preference for food high in ethanol was probably not problematic because ethanol was not available in large quantities. In contrast, someone who consistently desires large amounts of ethanol today can probably acquire an endless supply (although this may depend on how much money they have). This difference between modern and ancient environments is important in determining how and why mechanisms evolved and in considering how these mechanisms may come to play in modern times.

One particularly important brain mechanism which is often mentioned in biological explanations of drug addiction is the mesolimbic reward pathway. To fully understand the role of this pathway, it is necessary to first ask the question of why people use drugs. Essentially, they use drugs because drugs make them feel good. Further, the reason they feel good is because the drugs they use short-circuit neurological mechanisms that have evolved to play a key role in the manifestation of positive experiences. From an evolutionary point of view, the role of emotions is to motivate individuals to pursue adaptive experiences and avoid harmful experiences. Thus, adaptive activities such as eating and engaging in positive social interaction feel good; and this means that people are more likely to keep doing them. It has been theorized that the reason these sorts of activities feel good is because they activate the mesolimbic reward pathway in the brain. This pathway connects areas in the frontal cortex with the nucleus accumbens and the ventral tegmental area, in the limbic system. In terms of drug use, as suggested above, it is thought that when people use drugs they activate this pathway thereby creating positive sensations and experiences. The primary neurotransmitter that is widely believed to be instrumental in this pathway is dopamine and the effects of many drugs seem to be associated with changes in dopamine function. For example, cocaine and amphetamines block the reuptake of dopamine thereby increasing the levels of dopamine in the synapse.

Another important aspect of the biology of substance abuse and dependence is the research on genetic factors. Most of the research in this area has focused on the genetic component of alcohol addiction and there is now a growing body of research that shows that

alcoholism has a genetic component. Interestingly, what is most important in terms of the discussion herein is that genetic factors have been used to explain the varying levels of alcoholism across ethnic groups. The program of research that has explored this issue has focused on the genes that code for the liver enzyme aldehyde dehydrogenase (ALDH) which plays an important role in the metabolization of alcohol. Research indicates that the presence of ALDH2*2 is associated with a deficiency in ALDH2 which seems to protect individuals from the development of alcohol problems. Those who have this deficiency experience a range of aversive physiological consequences when they drink, due to a buildup of acetaldehyde. The presence of the ALDH2*2 allele has been found to vary across ethnic groups; it is close to zero in Europeans, Africans, and Native Americans, but it is present in between 30 and 50% of Northeast Asian populations.

A number of researchers have reported finding an association between the presence of the ALDH2*2 allele and a reduced likelihood of alcohol abuse and dependence. Furthermore, the association has been found to be particularly strong in individuals who carry two copies of the ALDH2*2 allele. These individuals have been found to have almost no risk of developing alcohol dependence. However, it is important not to overemphasize the role of genetic factors in addiction as research has also found that the role of genes is not absolute but is influenced by cultural factors. For example, a study which examined the protectiveness of the ALDH2*2 allele over time in Japanese populations found that the protective effects decreased as Japanese people became more accepting of alcohol and increased their consumption. Thus, it has been hypothesized that over time Japanese peoples' use of alcohol yielded to the social and cultural pressures that rewarded its use and paid less attention to the physiological effects.

So, why are evolutionary and biological factors important in understanding addiction across cultures? They are important because they are a part of the causal picture. The process of evolution and its resultant biology inevitably plays out in the manifestation of every drug use problem. While sociocultural factors are also important, neural processes provide the foundation for the thoughts, feelings, and behaviors of the individual drug user. Also, as illustrated above, there may be subtle differences across ethnic groups in the physiological effects of substances on the human body. Furthermore, in terms of substance abuse, and in particular, addiction, there are significant physiological changes that take place within the individual which contribute significantly to their ongoing drug dependence.

NORMS, VALUES, AND EXPECTATIONS

Social norms are widely believed to account for differences in substance use and abuse across cultures. The social norms surrounding the use of alcohol in various cultures provide a good example of this. There are many cultural groups which prohibit the consumption of alcohol in all contexts. These groups include Mormons, Muslims, Alcoholics Anonymous organizations, and some Protestant factions. However, it is important to keep in mind that teetotalism is not strictly adhered to by all members of these groups. For example, in some Muslim groups, consumption of alcohol is not unheard of.

Many cultures, while allowing the use of alcohol, nonetheless have norms which discourage excessive alcohol consumption and drunkenness. This is seen in particular in Jewish culture, but also in Chinese, Italian, and French cultures. However, these norms are not static; they change over time and they differ across various sectors of society. Also, the strength of the influence of cultural norms and values depends on the extent to which an individual is immersed in his or her cultural environment. Interestingly, one study found that Russian Jews who had immigrated to Israel had higher rates of alcohol use than Israeli Jews – perhaps because of the culturally accepted pattern of heavy drinking in Russia. Thus, in some cases, individuals who move to a new cultural environment will continue with the pattern of substance use that they learned in their country of origin. Another study found that American-born Chinese generally consume more alcohol than Asian-born Chinese and this exemplifies the significance of social–cultural factors in patterns of use.

One of the most widespread cross-cultural norms pertaining to the use of alcohol relates to women. In almost all countries, alcohol use is seen as being less acceptable in women than it is in men. Furthermore, in some countries it is forbidden for women to drink. In Nigeria and Lesotho, women who abuse or become dependent on alcohol are judged harshly by their society and are often marginalized and ostracized. In other societies, such as Japan, Mexico, and Peru, the stigmatization is not as extreme but still drinking in women is widely discouraged.

In many Western societies, heavy drinking by both men and women is viewed as normal. Although heavier patterns of drinking are more often seen in young men. In New Zealand, individuals who refuse to drink in social settings may be subjected to criticism and ridicule by others. Thus, there is often social pressure to binge drink to the point of significant physical illness. Also, in many Western societies, even the signs of physical illness (e.g., vomiting and loss of consciousness) that

accompany binge drinking are normalized. However, in contrast, alcohol addiction would probably not be seen as normative in most Western societies and would normally be frowned upon by mainstream society. While regular excessive alcohol consumption is seen as reasonably normal and acceptable, the sort of use seen with dependence would not be accepted. On the other hand, dependence on other drugs, such as methamphetamine, would probably be normalized in some social groups in the West. Therefore, whether addiction is seen as acceptable in Western societies would depend on the drug and the social context.

Another important dimension of the interaction between cultural factors and the effects of a substance on an individual is the beliefs and expectations of the individual. In Mexico, drinking sessions associated with drunkenness and violence are common. Similar links between alcohol and violence are also seen in some Papua New Guinean tribes. And of course, in many Western countries, such as the United States and New Zealand, there is a strong association between drinking and violence. However, in contrast, there are many cultural environments in which high levels of alcohol consumption are not associated with violence, such as among the Camba of Eastern Bolivia and the Yaruna Indians who reside in the Amazon. Thus, many researchers have concluded that the behaviors that arise following the consumption of alcohol depend on the beliefs and expectations of the individual and these have their roots in cultural norms.

Several studies have explored the connection between alcohol use and expectations (often referred to as expectancies). They have found that expectancies have a profound effect on the individual's response to drinking. For example, studies have found that if people believe they have been consuming alcohol, when in fact they have been imbibing a nonalcoholic beverage, then they will behave as if they are intoxicated. Further, with regard to other substances, it has been reported that the cultural context influences the content of the drug user's experience. For instance, if an individual has a vision while taking a drug in the context of a ceremony, the vision will typically reflect the purpose or context of the ceremony.

Beliefs and expectations can also influence the process of addiction. For example, some researchers have argued that if people believe that they have no control over their drug use, then they will be more likely to develop an addiction. Accordingly, if people believe that cocaine is highly addictive then they will be more likely to become addicted and to have an ongoing addiction problem. The drug-related attitudes and beliefs that an individual adopts probably come from his or her immediate sociocultural environment. Particular important sources within that environment might include media reports, cultural norms, personal experiences, and family knowledge that is passed down over the generations.

It is useful to consider the way that tobacco-smoking habits have changed in many Western countries over the last few decades. While rates of smoking have been falling steadily in many countries over the last 50 years, there has been a more significant drop-off in the last two decades. This is mostly explained in relation to the increased understanding of the harmful effects of tobacco on human health. However, this message has been delivered to varying extents around the globe. While some countries require that health warnings are placed prominently on cigarette packets, some countries do not. Also, rates of cigarette use vary around the world and it remains popular in many Asian countries. What is perhaps most important in regard to this discussion, is that the social norms surrounding the use of tobacco in the West have changed markedly such that it is no longer considered to be socially acceptable. Subsequently, many people in Western countries now hide their addiction to tobacco. In this way, the attitudes, beliefs, and behaviors surrounding tobacco addiction have changed over time and these changes have been associated with a significant reduction in tobacco use.

RECONCEPTUALIZING DRUG ABUSE AND ADDICTION: A CULTURE-INCLUSIVE APPROACH

Over the years, there have been many different approaches to classifying mental disorders and various mental disorders have come and gone from taxonomic systems. For example, diagnostic manuals previously included disorders such as drapetomania (which referred to slaves who ran away from their masters) and childhood masturbation disorder. What is most telling about the appearance and subsequent disappearance of these sorts of disorders is that the perception of what constitutes a disorder has changed. Furthermore, in the case of drapetomania, the social context has changed, rendering the so-called disorder redundant and irrelevant. Of course, looking back, it would seem that this disorder was borne out of a desire for social control of a particular group of people. So, to some extent mental disorders are social constructions which are influenced in important ways by the social and cultural and social context in which they are utilized.

Understandings of addiction have also risen and fallen on the tide of social change. At times, the disease model of alcoholism has been especially popular and at other times, social and psychological models have

prospered. Anyone familiar with the disease model debate will know that it continues to this day, although perhaps it is reasonable to suggest that there is now a growing body of researchers who take the view that alcoholism is caused by a combination of biological and social factors. With regard to drug problems, the conceptualization of the phenomenon is important because it is typically closely tied to the treatment approach. For example, generally those who adhere to the disease model of alcoholism recommend that alcoholics should abstain from using alcohol because the disease renders them helpless to control their use even if they attempt to consume very small quantities. This is of course the approach of Alcoholics Anonymous.

Arguably, one of the most influential contemporary models of mental disorder is unable to take into account the cross-cultural diversity of addiction. The model espoused by the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (DSM) is based on a biomedical view. Accordingly, discrete diagnostic entities are listed, each of which is associated with a particular treatment approach. There are four key aspects to the DSM definition of mental disorder: distress, disability, expectability, and dysfunction. The notion of distress captures the subjective and experiential aspects of mental disorder while disability refers to the more observable sequelae. The concept of expectability is understood with reference to statistical norms; so a condition is seen as abnormal if it is unexpected in a particular context. It is in this aspect that the DSM takes context into account. Finally, the term dysfunction, which may be "behavioral, psychological, or biological," encapsulates the idea that there is a breakdown in the normal function of the person.

In terms of taking account of cross-cultural differences in addiction, the DSM definition is problematic. One key problem with the definition is that in some groups there may be no distress associated with the drug dependence. For example, people who are on methadone maintenance typically have an addiction to an opiate; however, they may not be distressed by it. Similarly, tobacco is used all over the world and is highly addictive; however, many people are not distressed by their addiction to the substance. Similarly, at certain times and places in history, addiction to substances such as coca, opium, and coffee has not necessarily been distressing for people. Coffee is a good example because it is a commonly used addictive substance that is socially condoned and its use is not usually associated with feelings of distress.

Another problem with the DSM definition of mental disorder lies in the use of the term expectability. As outlined above, this concept is used to stipulate that in order for a condition to meet the criteria for a mental disorder it must not be expectable (or statistically likely) within the particular cultural group in which it occurs. With regard to drug addiction, this is problematic because drug use and even dependence is socially accepted and reasonably common in some populations. For example, inhalant use and addiction among young aboriginal Australians is so common in some settings that it is not unexpected or unlikely. Similarly, opium dependence among some Afghanistan communities has reached very high levels in recent years; in some areas, it is estimated that approximately half of the community is addicted to opium and in some cases whole families, including children, are dependent on the substance.

In response to the lack of cross-cultural applicability of the DSM model of mental disorder, the writer developed an alternative model that attempted to take account of the diverse nature of mental disorders. Based on a constructivist approach, the model aimed to provide a more flexible and culturally sensitive understanding of mental health problems. Constructivism is a philosophical approach that lies on a continuum between social constructionism and realism. It does not deny the possibility of objective knowledge; however, it sees subjective experience and interpretation of the empirical world as important and capable of influencing the process of knowledge acquisition. According to the constructivist model of mental disorder, there are four key components, all of which interact with each other. As shown in the model below, these components are cultural–historical variables, biological variables, psychosocial variables, and the self (Fig. 5.1).

Cultural–historical variables are included in the model in order to acknowledge that drug-taking practices and addiction often arise out of historical and cultural traditions. The forces of history play out in the way that individuals understand a drug and in the way that they interact with it. For example, the way in which kava is used in various Pacific Island settings today owes much to the way it was used in decades past in Pacific societies. As people grow up in various societies, they imbibe the attitudes and beliefs of previous generations. Furthermore, current legislation may also impact on drug use and dependence and this too has evolved over time.

As the name suggests, psychosocial variables incorporate both psychological processes and social factors. This variable rests on the assumption that these two factors have an important interaction. The individual's psychology is embedded in a social context and cannot be understood outside of this context. An example of the significance of this variable can be seen in New Zealand where there is a significant methamphetamine dependence problem. Generally speaking, individuals are introduced to this drug by their social group and, furthermore, it is often normalized within this group.

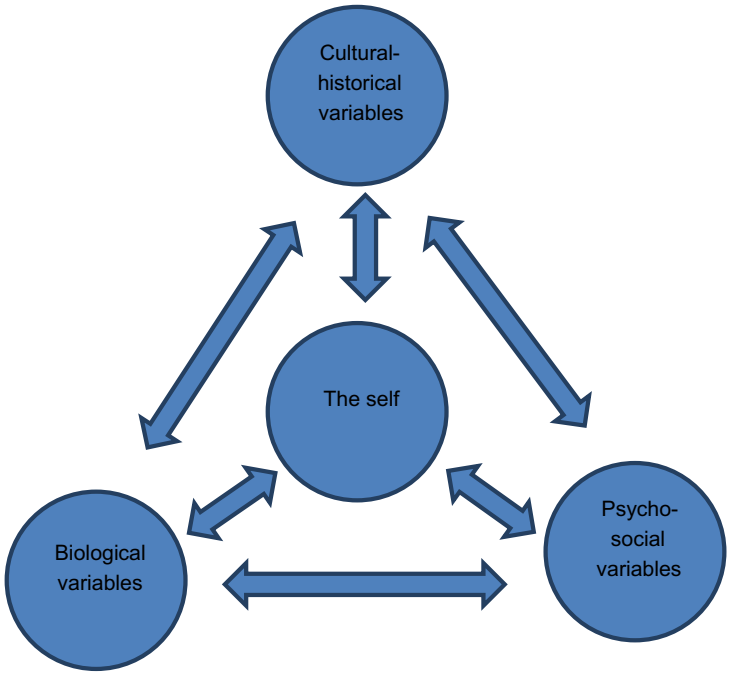


FIGURE 5.1 A constructivist definition of mental disorder. *Adapted from Thakker, J., Ward, T. & Strongman, K.T. (1999). Mental disorder and cross-cultural psychology: A constructivist perspective. Clinical Psychology Review, 19, 843–874.*

Thus, people begin using methamphetamine because their peers are using it and making it available to them. Inhalant use in aboriginal Australians is also an example of the importance of sociocultural factors as inhalant use is normalized in some aboriginal populations.

As explained earlier, biology is also important in explaining drug use and dependence. Drug addiction occurs because of the presence of particular physiological mechanisms. Furthermore, drug dependence involves a range of physical phenomena, in particular, withdrawal symptoms on attempting to stop the drug use and tolerance with chronic use. Also, different drugs have different main effects and side effects which may also influence the experience of drug dependence.

The final part of the model, the self, is the central part of the model, through which each of the other three factors interact; although they may also interact directly with each other. The self is conceptualized as being central to the model because regardless of the level of influence of the other factors, the individual is always unique and his or her drug problem is therefore also unique. The individual will inevitably interpret and respond to social, cultural, and historical influences and he or she will not take on such influences without processing them. Also, with regard to biological factors, while these will be similar across individuals, people will nonetheless respond to these factors in unique ways. Two people exposed to the same drug in the same setting may respond to the presence of the drug in different ways. One may experiment with the drug while another may not. One may become addicted to the substance while another may use it several times and thereafter desist in using it.

One important aspect of this model is that it is flexible. In explaining mental disorders, different factors may be more important for different disorders. For instance, biological factors are particularly important in explaining dementia whereas bulimia nervosa is probably best explained in terms of social and cultural factors. Similarly, in terms of conceptualizing drug dependence, different factors may be more or less salient depending on what is being explained. For example, if we want to understand why regular use of alcohol leads to tolerance, it is probably best to refer to biological theories. However, individuals may experience tolerance to alcohol in different ways depending on their beliefs and expectations and on the cultural context in which they are drinking. Thus, psychological and social explanations will also be important in understanding the process of tolerance.

The strength of this model lies in the fact that different parts of the model can be more or less salient, depending on what the investigator is attempting to explain. Also, it takes account of both historical and cultural variables and thereby acknowledges the broader context of the addictive process. As highlighted in this discussion, drug use and drug addiction differ around the world; therefore, any model of the addictive process needs to be sufficiently flexible to be able to take such differences into account.

DIVERSE TREATMENT APPROACHES

There is a growing understanding among clinicians and researchers that in order to tackle substance abuse and dependence, cultural factors need to be acknowledged and incorporated into the treatment process. As suggested above, an individual's drug-taking experience will be influenced by his or her attitudes and expectations, and by the norms and practices of his or her sociocultural environment; thus, it is important to consider such aspects in treatment.

First, it is important to acknowledge that there are differences across cultures in the availability and accessibility of treatment. There may be pragmatic barriers, such as cost, which may be more problematic for ethnic minorities, especially if they are more likely to be in lower socioeconomic groups. Also, there may be differences in the way that the treatment is perceived by various cultural groups. For example, in Japanese society, psychological problems are heavily stigmatized and are often believed to bring shame on an entire family. Therefore, individuals in Japan may avoid treatment as engaging in treatment would signify that the individual had a problem. Also, it has been reported that some African-Americans and Hispanics in the United States are reluctant to engage in methadone maintenance because they believe it is a form of control and oppression.

Around the world a number of culture-sensitive treatment programs have been developed. In New Zealand, specific treatment approaches for drug use problems have been designed for Maori. These approaches typically utilize indigenous models of health and wellness which tend to take a more holistic view in responding to health problems. One such model is "Te Whare Tapa Wha" which translates as the four-sided house. The "Te Whare Tapa Wha" model uses the image of a house to depict the four key elements of well-being, namely "te taha tinana" (the physical body), "te taha hinengaro" (the psychological aspect), "te taha wairua" (the spiritual realm), and "te taha whanau" (the family and community). The image of the house is used to demonstrate the interaction and interdependence of all of these elements; for example, if one aspect is weak then the house itself is weakened. This is a particularly significant metaphor for Maori as traditionally, the house or "whare" is seen as the spiritual center of the community. Furthermore, the model highlights the importance of taking a holistic and integrated approach to health and well-being.

In treating substance abuse and dependence, there are a number of ways in which the model can be utilized. The model may be used to conceptualize the individual's drug use problem through an examination of how well he or she is functioning in each of the key areas. For example, an examination of whanau links may reveal that the individual is socially isolated and not well connected with family. On the other hand, an analysis of "tinana" may indicate that the person has an unhealthy lifestyle characterized by poor eating habits and little or no physical activity. The model may also be used as an overarching framework for the delivery of conventional treatment approaches. For instance, cognitive behavior therapy may be delivered via the "Te Whare Tapa Wha" conceptualizations of health and well-being. Applications of the model in treatment aim to encourage clients to develop a strong identification with their culture and with the predominant beliefs and ideas, which form the basis of their culture with the goal of strengthening their self-belief and their sense of empowerment.

Similar approaches to substance use problems have been taken in Canada in the provision of treatment to indigenous populations. Clinicians there are now beginning to understand the importance of incorporating indigenous understandings of human health in their treatment programs. Again, there is a focus on family and community, and on traditional concepts of spirituality. Like the approach in New Zealand, the emphasis is on seeing the person as a holistic being who is connected with his or her physical and social environment. Also, many of the Canadian programs foster a strong sense of cultural identity through reference to indigenous norms, values, and traditions.

One of the key aspects of these sorts of drug treatment approaches in indigenous populations is the acknowledgment that social and cultural factors play an important role in the development of the problem. In some cases, individuals are socially isolated and there is a breakdown in their family connections. Also, they may experience a range of significant life problems such as lack of education, limited income, and lack of employment opportunities. Therefore, treatment needs to take a broad approach and attempt to address the underlying factors that may contribute to a person's drug-related problems. Also, common among indigenous treatment approaches is the inclusion of a spiritual component. This aims to encourage drug users to contemplate issues of meaning and purpose in their lives. It is not necessarily related to organized religion, but rather to the idea that developing a strong sense of meaning in life will assist in establishing clear future goals and finding fulfillment. In this way, the drug then becomes less appealing and hopefully inconsistent with the new outlook.

CONCLUSION

Drug use can be traced back many thousands of years. For millennia, human beings in many different parts of the world have been using drugs for various purposes. Perhaps they were primarily used for medicinal purposes originally, but over time they began to be used for their mind-altering properties. It is difficult to ascertain when the problem of addiction first arose although it was probably not seen in preagricultural times. Then, as human beings began to settle in particular areas and grow crops, supplies of drugs would have become more stable and reliable. However, the loss of productivity that would have been associated with developing an addiction would have meant that those who developed such problems would most likely have been unpopular or perhaps even ostracized. So historically such problems were probably uncommon. Likewise, today addiction is generally not seen as being socially acceptable in any sociocultural groups.

Nonetheless, addiction is present in many populations around the world. Most likely, this ubiquity is due to the presence of specific pathways in the brain and a common pattern of physiological processes, although there are some biological differences across ethnic groups in the way that some substances are metabolized. However, in explaining substance abuse and addiction, it is important to examine the individual's unique perspective and his or her sociocultural context. Thus, any model of substance use problems needs to be multifaceted and able to incorporate different levels of explanation. Further, in responding to substance abuse and addiction, it is essential to acknowledge and incorporate cultural-specific factors so that the individual is comfortable with the treatment approach. In a growing number of Western countries, including New Zealand and Canada, specific treatment approaches have been developed for indigenous people which utilize indigenous models of health and well-being and these have shown promising results to date.

SEE ALSO

An Evolutionary Perspective on Addiction, Historical Understandings of Addiction, Cultural Influences on Youth Alcohol and Drug Use

Glossary

Analgesic drug that is used to relieve pain.
Allele member of a pair or series of genes that occurs at a specific location on a specific chromosome.
DSM Diagnostic and Statistical Manual
Dioscorides Greek physician and botanist who lived from 40 to 90 AD.
Enzyme protein that catalyzes chemical reactions.
Ethanol pure alcohol.
Herodotus Greek historian who lived from 484 to 425 BCE.
Limbic system set of structures in the human brain which is located between the brain stem and the inner cortex. It influences the endocrine system and the autonomic nervous system and is involved in a variety of functions, including emotion and behavior.
Materia medica a multivolume work by Dioscorides which documented the use of plants, animals, and minerals for medicinal purposes. Written between 50 and 70 AD, it remained in use for approximately 1500 years.
Mesopotamia an area surrounding the Tigris-Euphrates river system which covers modern-day Iraq, Turkey, Syria, and Iran.
Neolithic also referred to as the New Stone era, this is the latter part of the Stone Age, when it is believed that *Homo sapiens* began using agriculture and domesticating animals.
Nucleus accumbens the main part of the ventral striatum in the human brain. It is understood to play a vital role in the experience of pleasure and reward.
Phylogeny the evolutionary development of a group of organisms.
Synapse a small gap, usually between nerve cells in the human brain, in which neurotransmitters convey nerve impulses from one neuron to another.
Ventral tegmental area a group of neurons located at the base of the midbrain. This region is understood to play a vital role in the experience of pleasure and reward.

Further Reading

Brodie, J.F., Redfield (Eds.), 2002. High Anxieties: Cultural Studies in Addiction. University of California Press, Berkeley, CA.
Durrant, R., Thakker, J., 2003. Substance Use and Abuse: Cultural and Historical Perspectives. Sage Publications, Thousand Oaks, CA.
Gazia, N., Connor, J.P., Ho, R., 2010. Cultural identity and peer influence as predictors of substance use among culturally diverse Australian adolescents. Journal of Early Adolescence 30, 345–368.
Prussing, E., 2008. Sobriety and its cultural politics: an ethnographer's perspective on "culturally appropriate" addiction services in Native North America. Ethos 36, 354–375.
Quintero, G.A., Lilliot, E., Willging, C., 2007. Substance abuse treatment provider views of "culture": implications for behavioural health care in rural settings. Qualitative Health Research 17, 1256–1267.
Vasquez, M.J., 2009. Latino/a culture and substance abuse. Journal of Ethnicity in Substance Abuse 8, 301–313.
Withy, K.M., Lee, W., Renger, R.F., 2007. A practical framework for evaluating a culturally tailored adolescent substance abuse treatment programme in Molokai, Hawaii. Ethnicity and Health 12, 483–496.

Relevant Websites

www.recoveryview.com – An online journal for professionals in the addiction and behavioral health fields.
www.ethnicity.de – Drug Treatment & Ethnicity – Update.
www.health.harvard.edu – Harvard Health Publications, a division of Harvard Medical School.
www.drugabuse.gov – National Institute on Drug Abuse.
www.alcoholrehab.com – Top Drug and Alcohol Rehab in Thailand.

CHAPTER

6

The Biopsychosocial Model of Addiction

*Monica C. Skewes**, *Vivian M. Gonzalez*[$]

[*]University of Alaska Fairbanks, Fairbanks, AK, USA, [$]University of Alaska Anchorage, Anchorage, AK, USA

OUTLINE

BIOPSYCHOSOCIAL VS. BIOMEDICAL MODELS OF ADDICTION

The biopsychosocial model of addiction posits that biological/genetic, psychological, and sociocultural factors contribute to substance use and all must be taken into consideration in prevention and treatment efforts. This model emerged in response to criticisms of the biomedical model, which has historically dominated the field of addiction studies. The traditional biomedical model was developed and is espoused by medical scientists for the study of disease, and its proponents also view addiction as a chronically relapsing brain disease with a genetic/biochemical cause. The biomedical or disease model of addiction views addiction as the manifestation of disturbances in measurable biochemical or neurophysiological processes in the afflicted individual. Contemporary medical disease models acknowledge the influence of social, psychological, and behavioral dimensions of addiction; however, these dimensions are viewed as relatively less important in the etiology and treatment of addiction. The medical disease model favors reductionism, whereby underlying biomedical causes for addiction are primarily implicated in the etiology/cause of the disorder, and mind–body dualism, where the mind and the body are viewed as separate and as not significantly affecting one another. Despite widespread favor among many scientists and healthcare practitioners, evidence from research studies of addictive behaviors does not support the medical disease model of addiction; instead, a biopsychosocial model that gives equal importance to biological/genetic, psychological, and sociocultural factors better fits the available data.

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00006-1

In 1977, psychiatrist George Engel authored a seminal paper calling for the abandonment of the biomedical model of illness in favor of a biopsychosocial model. Engel identified numerous problems with the biomedical model that would be alleviated by the adoption of a biopsychosocial model that recognizes biological, psychological, social, and cultural influences on illness. For example, the biomedical model views biochemical abnormalities as the cause of any illness, and posits that correcting the biochemical abnormality will cure the illness. However, in many disorders, a person may remain ill after the biochemical abnormality has been corrected and, conversely, a person may never become ill even in the presence of an abnormality. For example, when infected with the virus that causes the common cold, some research participants become ill and some do not. The biomedical model does not account for the finding that, among people with similar genetic predispositions or physiological problems, some people develop an illness while others remain well. Engel surmised that psychological and sociocultural factors must explain the differences in the disease state among people with the same biochemical abnormalities.

It has been well established that illness is not merely the result of biochemical dysfunction or abnormality, as some people become ill in the absence of an abnormality or dysfunction. The effects of stress on illness have been well supported in the literature, as has the role of expectation on illness and health. The placebo effect, where an inert ingredient can result in biochemical reactions for the person who believes he or she is ingesting a drug, is evidence for the role of expectation in illness, and supports Engel's view of a connected mind–body experience. There also is evidence for the importance of the patient–provider relationship in healing; if psychosocial variables were not important, it would not make sense for rapport building and communication between the physician and the patient to have such strong influences on health outcomes. Moreover, if illness is caused only by the existence of a physical abnormality, then it should be cured by correcting the deviance, but this is not always the case. Most illnesses, disorders, and syndromes, including disorders of addiction, are caused by the interaction of numerous factors – biological, psychological, social, cultural, cognitive, and environmental. Therefore, these factors must be addressed in order to result in a recovered state.

CONCEPTUAL MODELS OF ADDICTIVE BEHAVIOR

A discussion of helping and coping by Brickman and colleagues identified four models of addiction based on beliefs about attributions of responsibility for acquiring the addictive problem and the responsibility for solving the addictive problem. The moral model holds that people who suffer from problems of addiction are responsible for both acquiring and solving the problem. People who become addicted are seen as morally weak with poor willpower, and they must will their way through addiction in order to recover. There is little support for this model in the literature. The enlightenment model holds that the person is responsible for developing the addiction, but is not responsible for solving the problem. The enlightenment model is espoused by Alcoholics Anonymous and other 12-step philosophies, and requires people to seek recovery by turning the problem over to a higher power. Only a higher power can cure addiction, and it is the person's task to form and strengthen a relationship with a spiritual entity so that this entity can solve the addiction problem. The medical/disease model emerged in response to the moral and enlightenment models that placed blame on the addict for his or her problem. In the medical model, the addict is responsible neither for the development of the problem nor for its resolution. This model posits a biological/genetic predisposition for addiction, an underlying disease process, and assumes that the disease is progressive. The medical/disease model fails to account for the finding that many people with problems of addiction do recover without professional treatment. Finally, the compensatory model holds that people are not responsible for developing the addictive problem, but are responsible for their own recovery. In the compensatory model, the role of multiple factors in the development of addictive behavior is noted (including biological predisposition, early experiences, and social and cultural variables), and the continued use of substances is viewed as a way to cope with stress. Of these four models, the compensatory model is the most similar to the biopsychosocial model.

BIOPSYCHOSOCIAL MODEL OF ADDICTION

Science has not discovered a single factor that can explain why some people are able to use substances without progressing to addiction, while others abuse or become dependent on substances. Instead, the available evidence suggests that biological, genetic, personality, psychological, cognitive, social, cultural, and environmental factors interact to produce the substance use disorder, and multiple factors must be addressed in prevention and treatment programs. The interaction of these factors to produce substance use problems is the core tenet of the biopsychosocial model of addiction. This model is a way to understand and explain the problem of addiction, but has not generated testable

hypotheses as have theories of behavior change like the Health Belief Model or the Theory of Reasoned Action/Theory of Planned Behavior (TRA/TPB). The essence of the model is that the mind and the body are connected and both the mind and the body affect the development and the progression of addiction within a social and cultural context. Only by considering all of these factors can addiction be accurately conceptualized.

BIOLOGICAL FACTORS AND THE DEVELOPMENT OF ADDICTIVE BEHAVIORS

Given the right environment, biological and genetic predispositions may increase the risk of substance use problems. Adoption and twin studies have found that substance abuse is to some extent heritable. Male children of an alcohol-dependent parent have four times the risk of becoming problem drinkers compared with the children of nondependent parents, while female children of alcohol-dependent mothers evidence a threefold greater risk. It has been reported that 30.8% of people with alcohol dependence had at least one alcohol-dependent parent. Among adults with alcohol dependence, 27% have alcohol-dependent fathers and 4.9% have alcohol-dependent mothers, compared with alcohol dependence among 5.2% of fathers and 1.2% of mothers of people without alcohol dependence. Among twin pairs in which one twin was diagnosed with alcohol dependence, there is a significant difference in the proband concordance rate among monozygotic (54.2%) and dizyogtic twins (31.5%). Calculated heritability ranges from 40–90% across studies, with more chronic and severe forms of alcohol dependence showing greater estimates of heritability. However, it is important to note that someone with a strong genetic predisposition to addiction still needs to engage in substance use before the addictive behavior becomes manifest.

Once alcohol is consumed, however, children of an alcohol-dependent parent experience the effects of alcohol differently than the children of nondependent parents. For example, research on subjective experiences of alcohol intoxication and body sway while intoxicated found that sons of an alcohol-dependent parent respond less intensely to moderate doses of alcohol. When given the same amount of alcohol as controls, sons of an alcohol-dependent parent had less body sway and were less likely to report feeling intoxicated. Follow-up studies have found that decreased subjective intoxication predicted later development of alcohol use disorders. Other studies have found that the children of an alcohol-dependent parent are less sensitive to the negative consequences of alcohol, resulting in increased alcohol consumption. Further, sons of an alcohol-dependent parent have decreased EEG alpha rhythms, also found in people with current alcohol dependence. Other studies have found that the sons of an alcohol-dependent parent have lower language functioning, lower learning achievement, lower verbal intelligence, and other neuropsychological differences when compared to controls. There is evidence that children of an alcohol-dependent parent who become alcohol dependent themselves have a worse prognosis than alcohol-dependent people who are not the children of alcohol-dependent parents. For example, the children of an alcohol-dependent parent show symptoms of alcohol problems earlier, have greater physical dependency on alcohol, and report less control over their drinking.

A genetic predisposition toward addiction does not influence the substance of choice to which a person may become addicted; instead, it is associated with an increased propensity toward addictive behavior in general. It also is important to note that genetic factors may be protective against alcohol use disorders. People of Asian descent are more likely to lack one isozyme of a liver enzyme known as alcohol dehydrogenase that aids in the metabolism of alcohol in the liver. People with this genetic variation have a flushing reaction to alcohol, characterized by flushed, reddish skin, and are much less likely to ever develop alcohol problems.

Research from the fields of genetics and biochemistry has identified other biological risk factors for addiction. People with impulse control disorders, including people with substance abuse problems and gamblers, are statistically more likely to have the dopamine D2A1 gene than controls. This genetic polymorphism is associated with reduced D2 receptor density and deficits in the dopaminergic reward pathway. Research has found that those with low D2 receptor density are more likely to seek out pleasurable activities including alcohol use, drug use, and gambling. This may translate into increased likelihood of experiencing problems associated with addictive behaviors.

Further evidence of the heritability of the risk for alcohol dependence can be found in animal studies. Researchers have been able to use selective breeding to develop strains of rats that differ in their liking of alcohol. One strain of rat (C57BL/6) has been bred to prefer alcohol over water. These animals seek out alcohol, ingest it willingly, engage in efforts to get alcohol, and become physically dependent on it, showing signs of tolerance and withdrawal. Other strains of rats have been bred to self-administer other drugs of abuse at high rates. The fact that an alcohol-preferring strain of rat has been developed is strong evidence of the influence of heritability on alcohol use behavior. Furthermore, studies have found deficits in serotonin in particular brain regions of rats that have

been bred to like alcohol. Despite the strong evidence of the role of genetic influence on alcohol use behavior, biology is still insufficient to account for the entirety of the problem. There still remain cases where people with no known genetic risk become addicted and cases where people with great genetic risk do not. The biopsychosocial model of addiction acknowledges that psychosocial variables also are needed to explain these occurrences and that these variables may interact with genetic and biological risks to cause addiction.

PSYCHOSOCIAL FACTORS AND THE DEVELOPMENT OF ADDICTIVE BEHAVIORS

Researchers have discovered consistent predictors of drug use initiation and subsequent use across multiple substances of abuse, including personality variables, learning factors, and higher-order cognitive processes. Substance abuse is highly comorbid with affective disorders and other psychiatric diagnoses, although some psychiatric problems (e.g. depression and anxiety) may be effects of the substance use as well as causal factors. Many (but not all) substance abusers have a history of antisocial behavior, nonconformity, deviance, acting out, impulsivity, and low self-esteem; however, these also can be the effects of substance misuse. Research establishing the role of psychosocial factors in the development of addictive behaviors provides evidence that addiction is a multifactorial problem, not a disease solely caused by a measureable underlying physiological abnormality or deficit, and provides support for the biopsychosocial model of addiction.

Risk Factors in Children

Much research has been conducted on childhood variables that increase the risk for alcohol dependence and substance use disorders. Consistently found in the literature is evidence for an increased likelihood of addiction among children who are victims of abuse and who exhibit externalizing behaviors such as those seen in conduct disorder, attention deficit/hyperactivity disorder, and oppositional defiance. In particular, antisocial and deviant behaviors such as aggression, hostility, vandalism, sadistic behavior, rebelliousness, and association with deviant peer groups place one at risk for substance use disorders later in life. One study found that problem drinkers exhibited more externalizing behaviors in childhood than did moderate drinkers, and moderate drinkers exhibited more of these behaviors than did light drinkers. Other research has found that tolerance of deviance in adolescence is a strong predictor of alcohol and other substance abuse in adulthood. Antisocial personality disorder is highly comorbid with substance abuse and dependence, and antisocial behaviors in childhood are strong predictors of substance problems in adulthood, independent of a family history of substance abuse.

Personality and Temperament

Addictive behaviors result from the interaction between genetic predisposition and psychosocial variables, including personality and temperament. Personality variables that impact later substance use include high novelty/sensation seeking, low harm avoidance, negative affectivity, and reward dependence. Other temperament variables that predict later substance problems are low attention capacity, high emotionality, low sociability, and impulsivity. A difficult temperament in childhood – defined as a high activity level, low flexibility, low task orientation, mood instability, and social withdrawal – has been shown to predict substance abuse in adolescence. One research study found that a difficult temperament in childhood was a stronger predictor of later alcohol dependence than a family history of alcohol dependence. Regarding the Big Five factors of personality (neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness), a family history of alcohol dependence is positively associated with openness to experience and negatively associated with agreeableness and conscientiousness. Unconventionality and deviant behavior are strong predictors of substance abuse across multiple research studies. In addition to increasing risk for substance use, temperament may influence adolescents' decisions when forming peer groups, which may then directly impact substance use. Adolescents who are more deviant and less conventional tend to select peers who also are more deviant and less conventional, further enhancing their risk for substance abuse.

Classical and Operant Conditioning

Classical and operant conditioning are learning processes that affect animal behavior, including addictive behaviors among humans. Classical conditioning works to establish a link between reflexive, involuntary behaviors and antecedent conditions, whereas operant conditioning concerns the modification of voluntary behavior in response to its consequences. In classical conditioning, an unconditioned stimulus (US) is paired with a conditioned stimulus (CS), resulting in a conditioned behavioral response (CR) to the conditioned stimulus. For example, Pavlov's dogs learned to associate the sound of a bell (the CS) with food (the US) to produce salivation (the CR). After several pairings of the bell with food, the bell itself became sufficiently linked

with food to produce salivation even in the absence of food. Among people engaging in addictive behaviors, an unconditioned stimulus (e.g., drug paraphernalia) can become paired with a conditioned stimulus (the drug) to produce a conditioned response (psychomotor stimulation). People (such as an addict's drug-using social network), places (such as locations where drugs are purchased or used), and things (such as drug paraphernalia, alcohol bottles, or substance-related words) are linked to the unconditioned stimulus (the substance) and take on the role of conditioned stimulus, evoking a conditioned response (e.g., craving). Encountering the conditioned stimuli associated with substance use (i.e., triggers) is a strong precipitant of relapse among people in recovery from substance use disorders.

In operant conditioning, behavior is reinforced via punishment, positive reinforcement (reward), or negative reinforcement (the removal of an adverse consequence). Reinforcement is any consequence that increases or decreases the likelihood that a behavior will be repeated. Among people engaging in addictive behaviors, operant conditioning affects the probability that the behavior will recur. For example, smoking behavior may be positively reinforced by pleasurable sensations caused by nicotine and simultaneously negatively reinforced by the elimination of nicotine cravings. All drugs of abuse act on the central nervous system and initially produce pleasant feelings and a hedonic state, but people differ in how reinforcing they find these feelings to be. People who enjoy the sensations produced by substance intoxication (i.e., find intoxication to be positively reinforcing) are more likely to use substances to the point of intoxication again in the future than are people who do not enjoy the feeling of intoxication. One study found that the degree of perceived reinforcement following initiation of drug use was predictive of the magnitude of the resultant drug problem.

Classical and operant conditioning work together to produce a behavior chain, or a sequence of behavior that can be understood in terms of both its antecedents (classical conditioning) and consequences (operant conditioning). Antecedents are also known as cues. Once a behavior chain has been activated, each cue serves as the reinforcer of the behavior that occurred previously as well as the antecedent of the behavior that follows. Cued habitual behaviors are both classically conditioned and reinforced or punished via operant conditioning. For example, encountering a liquor store may serve as a classically conditioned cue for a problem drinker, which results in craving. Craving may then serve as a cue to consume alcohol, and this behavior may then be negatively reinforced by alleviating stress or negative mood. By determining and understanding the behavior chain involved in addictive behaviors, intervention can be aimed at breaking the classically conditioned link between a cue and the behavior, by altering the reinforcement for the behavior, or both.

Substance use functions as positive reinforcement when the pleasant effects of intoxication are interpreted as rewarding to an individual. At the same time, people use substances to cope with unpleasant emotions, to manage stress, and to alleviate negative symptoms of withdrawal. In these ways, substance use also functions as a negative reinforcement. The more frequently one uses substances as a reward or as a way of coping with negative emotions or life events, the stronger the association becomes and the more difficult it is to extinguish the substance use behavior. This partially accounts for the finding that treatment is more difficult and relapse is more likely among individuals who have longer histories of substance use. It also has been suggested that use of substances to cope leads to an erosion of alternative coping behaviors, thereby making continued substance use and dependence more likely.

Outcome Expectancies

One area of research that has uncovered some of the strongest and most reliable effects of psychology on addictive behavior is that of alcohol outcome expectancies. Addiction is not merely a physiological response to something that feels good and is rewarding; it is influenced strongly by the labeling, interpretation, and meaning that a person ascribes to a substance of abuse. Outcome expectancies are conditioned cognitions; this refers to a person's beliefs about the effects that using alcohol (or another substance) will bring about. People who develop substance use problems report that using a substance results in positive, desired effects such as the ability to avoid or escape negative mood states. Common alcohol expectancies, as identified and described by Alan Marlatt, include relaxation and tension reduction, positive global changes in experience, sexual enhancement, social and physical pleasure, increased assertiveness, and increased arousal and interpersonal power. People may learn what they can expect from alcohol from prior experience or vicariously; indeed, evidence suggests that one need not have experience with alcohol in order to form strong expectancies about its effects. Watching others model the behaviors associated with intoxication (e.g., becoming louder, becoming more socially confident and engaged, and developing looseness of speech) can teach an observer what the effects of alcohol consumption are, thereby creating outcome expectancies. Experience with drinking may then reinforce previously held beliefs about the positive effects of alcohol. Expectancies also influence motives to drink – people who state that they expect alcohol to help relieve tension are more likely to turn to alcohol when stressed. Heavier drinkers report more positive alcohol

outcome expectancies and fewer negative outcome expectancies than lighter drinkers.

Self-efficacy

Another psychological variable that influences the development of substance use disorders is Bandura's concept of self-efficacy. Self-efficacy is defined as an individual's belief in his or her ability to perform a certain behavior in order to achieve a desired outcome. Self-efficacy for substance use is developed when one observes a model obtain and use substances. For example, an adolescent may develop self-efficacy for smoking by observing peers purchasing cigarettes at a location that does not check identification, lighting a cigarette, and inhaling the smoke. The adolescent's confidence in his or her ability to smoke is thereby increased. However, self-efficacy also refers to one's belief that he or she is capable of handling a stressful or challenging situation without using substances. Research has found that people are more likely to use substances in situations where they feel unable to cope with the demands of the situation or negative affect. As one uses substances more and more often to cope with stress or other life problems, the use of other more adaptive coping strategies decreases, which then results in reduced self-efficacy for the use of these alternative coping skills. This also translates into decreased self-efficacy in one's ability to refuse substances in the face of challenging life circumstances.

Social Influences on Substance Use

Families

In addition to genetic factors, addictive behaviors are transmitted between generations in families due to social influences. Social Learning Theory posits that modeling influences behavior, and that adolescents who observe substance use in their parents are more likely to use substances themselves. This assertion is supported in the research literature. However, there is evidence that modeling is not the only way in which parental influence on adolescent behavior takes place – parents also influence adolescents' behavior via norms and perceived attitudes. Numerous studies have found support for the association between parental approval of substance use and adolescent use of alcohol, tobacco, and marijuana – adolescents whose parents have positive attitudes toward substance use are more likely to use substances. Among college students, perceptions of parental approval of alcohol consumption were positively associated with experiencing a drinking problem. Perceived parental approval of illicit drugs was found to predict earlier first use of drugs and increased current frequency of drug use. Families also play a protective role against the development of substance abuse. Parental monitoring (supervision) and consistent discipline are associated with lower risk for substance abuse among children. Among women, becoming a parent also is associated with decreased risk for drinking problems.

Peers

Peers influence adolescents' values, attitudes, and behavior in multiple domains, including substance abuse. Having a peer group that uses substances is a strong predictor of adolescent substance use, as is the perception that one's peer group endorses substance use. When adolescents associate with peers who hold socially deviant attitudes and beliefs, the risk of substance use increases. Friends' smoking is among the strongest predictors of adolescent smoking behavior. Peer group involvement is thought to impact substance use through interaction with other risk factors, including family problems, stress, mental health, and self-esteem. Among adolescents who drink, the most important reasons for alcohol use were to socialize with friends, cope with tension and anxiety (especially regarding interactions with the members of the opposite sex), improve mood, and alleviate boredom. Male adolescents, who have higher rates of alcohol use than females, also have higher rates of involvement with peer groups that maintain deviant attitudes.

Peer influences on substance use behavior are not only important during adolescence. Studies have shown that alcohol use among adults is likewise influenced by peer drinking. College students' alcohol use was found to be positively correlated with their friends' alcohol use and with the students' perceptions of their friends' drinking. Often, college students who drink heavily report that their peers drink at the same levels as they do. Interventions for college students in which they are given feedback about how much they are drinking in relation to normative drinking for peers of the same gender and age demonstrate that high-risk drinkers are in fact consuming more alcohol than is normative for their peer group. Normative feedback interventions also demonstrate that heavy drinkers overestimate what is normative drinking, such that they erroneously believe most students drink as they do. These interventions consistently have been shown to result in decreased alcohol consumption and related problems for college students. Other research has found a positive correlation between alcohol use in adults and their perceptions of their peers' alcohol use. Heavy-drinking adults report having larger drinking social networks than do light or moderate drinkers. Social networks are important influences on adult substance use as well as adolescent substance use. Several studies have found support for the assertion that greater alcohol

involvement among one's peer network is associated with heavier drinking among both men and women. This relationship is independent of sociodemographic and individual difference variables and alcohol expectancies.

Spouses and Intimate Partners

Among adults, spouses and intimate partners are the most important and influential social connections people have. Research has supported spousal concordance in substance use behavior – wives and husbands tend to use the same substances, and heavy drinking or drug use in one partner predicts heavy use in the other partner. Problem drinkers are disproportionately more likely to be married to other problem drinkers. Marriage, however, also is a protective factor as it is often associated with a decrease in drinking.

Other Individual Difference Variables that Influence Substance Use

Ethnicity and Culture

National surveys have documented racial/ethnic differences in rates of substance use, with certain ethnic minority groups (e.g. African Americans, American Indians/Alaska Natives) reporting disproportionately high rates of substance use and dependence. There are numerous reasons for disparate rates of use and dependence among ethnic minorities, including increased risk factors such as poverty, discrimination, microaggressions, and stress among minority groups. Ethnicity also has been shown to moderate the associations between the risk factors and substance use. One study found that substance availability and perceived parental approval differed among Whites, African Americans, and Asians, resulting in ethnic differences in substance use initiation and stated intentions to use substances. With regard to cigarette use, White adolescents were most strongly influenced by adult and peer smoking; for African Americans, however, risk taking was a stronger predictor of smoking behavior than social norms. Among Hispanics, the perception of adult and peer approval of smoking was the strongest predictor of adolescent tobacco use; among Asians, peer and family influence was not as important for the prediction of adolescent smoking as poor academic performance and low self-esteem.

Other research has found that ethnic labels and ethnic identity influence substance use. In a study of middle school students, having a strong sense of ethnic pride was protective for African American, Mexican American, and mixed-ethnicity students, as these students reported less exposure to drugs and less drug use than those with lower levels of ethnic pride. The opposite was true for ethnically proud White students, who reported greater drug exposure and use. Also, ethnic minority students who reported having behavior, speech, and looks that are common in their ethnic group reported greater drug exposure and use, whereas White students who viewed their behavior, speech, and looks as typical of their ethnic group reported lower exposure and use of drugs. Among Mexican Americans and American Indians, a stronger sense of ethnic identity was associated with decreased substance use.

Gender

Across cultures, males smoke more, use more drugs, consume more alcohol, and have more alcohol and other substance use disorders than females. Gender differences in alcohol consumption may be due to differences in availability of alcohol and other substances; they also may be partially attributable to adherence to traditional gender roles. Females who espouse more traditional gender role attitudes drink less than their less conventional counterparts, while the opposite holds true for males. Among women, but not men, becoming a parent also is associated with decreased risk for drinking problems. Peer influence on smoking is stronger for females than for males. Research has found that females respond more than males to the reinforcing properties of substances and may develop drug abuse and dependence more rapidly than males. Other risk factors for addiction, such as childhood physical and sexual abuse, depression, intimate partner violence, and posttraumatic stress disorder may play a more important role in the initiation and maintenance of drug use among women than among men. One study found that women's tobacco use was more influenced by social factors than physiological dependence when compared with men's tobacco use. Women also report a greater propensity to use substances to alleviate negative affect than men.

Environmental Influences on Substance Use

Availability of substances has been shown to be a major factor in the initiation of substance use and the development of substance use disorders. In order for a genetic predisposition to result in addictive behavior, one must interact with the agent of addiction (e.g., alcohol or drugs). Increased availability of alcohol or drugs makes increased contact with substances possible. Rates of smoking are increased in areas with no smoking area restrictions and are decreased when laws prohibiting indoor smoking are enforced. The prevalence of alcohol use and alcohol use disorders is greater in neighborhoods with more bars and liquor stores, and research suggests that there are more bars and liquor stores in ethnic minority neighborhoods, which may account for some ethnic group differences in rates of alcohol-related problems. Socioeconomic

status also affects substance use behavior. Poverty is an established risk factor for alcohol and drug problems, and low income is associated with alcohol dependence and comorbid psychiatric disorders. Research on neighborhood disorganization, operationalized as high population density, physical deterioration, high crime rates, and the presence of illegal drug trafficking, has found that residing in disorganized neighborhoods strongly increases the likelihood of adolescent substance use.

Social engineering and public policy approaches to substance-related problems have had some success, adding further support to the biopsychosocial model of addiction. For example, when alcohol was made illegal during prohibition, national rates of drinking and associated health consequences such as cirrhosis of the liver were reduced. There also are areas within the United States (e.g., villages in rural Alaska) where alcohol is illegal. For example, the Alaska State Local Option Law, implemented in 1981, allowed Alaska Native communities to decide whether to permit drinking and what kind of alcohol control policies to implement. Research shows that the rates of homicide and accidental death are lower in communities that opted to ban the sale and importation of alcohol. Laws regulating the sale of alcohol and tobacco to persons under the minimum age limit have served to decrease use of these substances. Between 1970 and 1975, 29 states lowered the minimum drinking age with negative effects on public health; rates of adolescent alcohol consumption, alcohol-related injuries, and automobile fatalities increased during this period. When the minimum drinking age was later increased, there were reductions in adolescent alcohol consumption, alcohol-related injuries, and automobile fatalities. Raising the minimum drinking age also resulted in decreased prevalence of alcohol abuse and dependence in the following years. Research on taxation of alcohol and tobacco supports the use of taxes as a means to decrease alcohol and tobacco use. Among 18–20 year olds, increasing the price of alcohol resulted in decreased automobile fatalities. Such approaches to ameliorating addiction problems would not make sense and would not be effective if the biomedical model of addiction was accurate.

THE BIOPSYCHOSOCIAL MODEL AND ADDICTION TREATMENT

The most successful addiction treatment programs incorporate strategies to enhance coping, reduce craving, manage triggers, and prevent relapse. Some programs involve medication, but pharmacotherapy often is not considered an essential part of recovery from problems of addiction. Instead, finding ways of managing difficult emotions, coping with negative life circumstances, enhancing social support for sobriety, and establishing a lifestyle free of substance abuse is essential to long-term recovery. Recovery from addiction requires a biopsychosocial approach with attention paid to biological, psychological, and social aspects of addiction.

Natural Recovery

According to the biomedical model, addiction is a chronically relapsing brain disease that will progress unless treated. The biopsychosocial model, on the other hand, recognizes the occurrence of natural recovery. Natural recovery, or recovery in the absence of professional treatment, is one of the most common methods of recovering from substance abuse problems. Studies of the reasons for natural recovery have identified many precipitants to change, including a meaningful religious or spiritual experience, suffering a loss, support from family and friends, a personal injury or illness, or the substance-related injury or the illness of another person. These reasons are external motivators, aspects of one's social and cultural environment. In recovering from addiction, strategies that support successful self-change include the establishment of social support for sobriety, adoption of new coping skills and stress management techniques, and overall lifestyle changes such as changing one's social network, restructuring leisure time, and avoiding triggers that cause craving. One research study found that people with higher levels of motivation to change, greater commitment to change, more frequent and persistent use of coping strategies, and more frequent use of self-reinforcement strategies were more likely to be successful in self-change of addictive behaviors.

Medication

Medication is the treatment of choice for disease under the biomedical model. There are several medications that are helpful during substance abuse treatment, but no medication has been shown to cure or prevent addiction, as would be expected under the biomedical model. For smoking, nicotine replacement therapies such as the nicotine patch, gum, and nasal spray have been shown to double one's chances of successfully quitting. These tools work by gradually reducing the smoker's dependence on nicotine while extinguishing the smoking behavior, with the goal of slowly tapering off the nicotine until one is free of the substance. For opiate dependence, methadone and buprenorphine have been shown to be effective at reducing cravings and avoiding withdrawal symptoms, but these medications have side effects and other consequences. People in treatment for alcohol dependence may be prescribed

Antabuse to classically condition them to develop an aversion to alcohol, but this is only modestly effective under ideal circumstances. Naltrexone, an optiate receptor agonist, when combined with medical management, showed evidence of being equally effective in treating alcohol dependence as an intervention that combined elements of several effective behavioral treatments. Treatment for substance dependence in general may benefit from anxiolytic or antidepressant pharmacotherapy to help manage the negative mood that previously was managed by alcohol or other substance use. Overall, medication can be a helpful adjunct to psychosocial treatment, but no medication has been shown to solve the problem of addiction, and recovery in the absence of pharmacological treatment is common.

PSYCHOSOCIAL FACTORS IN THE TREATMENT OF ADDICTION

Readiness to Change

Of particular relevance to the treatment of addiction is the psychological factor of motivation or readiness to change. Research has found that motivation increases the likelihood that someone will seek out treatment and complete a treatment program or will be successful in changing his or her substance use in the absence of formal treatment. Greater motivation at treatment initiation also is predictive of better long-term outcomes. According to the Transtheoretical Model, a leading model of behavior change that was developed to explain how people change addictive behaviors, readiness to change is a function of decisional balance, or the cognitive appraisal of the pros and the cons of changing. People who rate the cons of changing their substance use as more salient than the pros of changing report being less motivated to reduce their substance use. People for whom the pros of changing outweigh the cons of changing their substance use are more motivated to change and have greater success in attempts to limit their consumption of substances. Readiness to change can be impacted by processes of change such as consciousness raising, dramatic relief, and self-liberation. Motivational interviewing is one therapeutic style that has been shown to be especially effective with substance users who are low in motivation to change and to bring about greater readiness to change. The extensive body of literature demonstrating the impact of readiness to change on subsequent substance use provides additional support for the biopsychosocial model over the biomedical/disease model of addictive behavior.

Self-efficacy

Self-efficacy is a cognitive factor that is crucial to recovery from addiction. In order to refrain from using substances, one needs to believe that he or she is capable of managing difficult situations in other ways. Self-efficacy can be built through experience and exposure to different behavioral options. For example, when a smoker foregoes a cigarette in favor of a brisk walk to manage stress, he or she builds self-efficacy by proving to himself or herself that a response other than smoking is possible for the individual, and that stress can be relieved in the absence of cigarette smoking. The greater experience one has with using coping strategies other than substance use, the greater his or her self-efficacy becomes.

SUMMARY

In summary, there is a great deal of evidence supporting the biopsychosocial model of addiction, which gives weight to biological, psychological, and social factors in understanding the development and progression of substance use problems. Research supports the role of biological factors such as genetic predisposition in the development of addictive behaviors. At the same time, psychological and cognitive factors such as outcome expectancies, self-efficacy, and readiness to change and social factors such as family, peer, and intimate partner influences on substance use are equally important. In order to prevent and treat addictive behaviors, attention must be paid to the biological, psychological, and social factors that interact to produce and maintain disorders of addiction. The evidence clearly indicates the importance of factors in all three realms, and successful treatment programs will benefit from taking a biopsychosocial view of the problem of addiction.

SEE ALSO

Behavioral Economic Factors in Addictive Processes, Cognitive Factors in Addictive Processes, Contextual Factors in Addiction, Disease Model, Families and Addiction, Gender Differences, Interpersonal Factors and Addictive Disorders, Natural Recovery, Peer Influences on Addiction, Personality and Addiction Processes, Craving and Expectancies

Glossary

Biopsychosocial the influence of biological, psychological (including cognitions, emotions, and behaviors), and social (including culture and environment) factors on the development, maintenance, and treatment of disease.

Health Belief Model the Health Belief Model is an empirically supported theory of health behavior change developed to understand and predict the use of health services. The model posits that health behaviors are influenced by perceived susceptibility to and perceived severity of disease, as well as perceived benefits and barriers to changing behavior. In recent years, two additional constructs (cues to action and self-efficacy) have been incorporated in the Health Belief Model to allow for the prediction of more general health behaviors.

Theory of Reasoned Action/Theory of Planned Behavior the Theory of Reasoned Action/Theory of Planned Behavior (TRA/TPB) is a combination of two psychological theories of health behavior change developed by Fishbein and Azjen to explain and predict human behavior. The models posit that the most important aspect of behavior change is one's intention to change. The intention comprises a person's attitude toward the behavior and subjective norms. Attitudes, in turn, represent a combination of behavioral beliefs, evaluations of behavioral outcome, normative beliefs, and the motivation to comply with recommendations. The model has empirical support and allows for the prediction of human behavior.

Further Reading

Borrell-Carrio, F., Suchman, A.L., Epstein, R.M., 2004. The biopsychosocial model 25 years later: principles, practice, and scientific inquiry. Annals of Family Medicine 2, 576–582.

Engel, G.L., 1977. The need for a new medical model: a challenge for biomedicine. Science 196, 129–136.

Engel, G.L., 1980. The clinical application of the biopsychosocial model. American Journal of Psychiatry 137, 535–544.

Engel, G.L., 1997. From biomedical to biopsychosocial: being specific in the human domain. Psychosomatics 38, 521–528.

Griffiths, M., 2005. A 'components' model of addiction within a biopsychosocial framework. Journal of Substance Abuse 10, 191–197.

Leonard, K.E., Blane, H.T., 1999. (Eds.) Psychological Theories of Drinking and Alcoholism. Guilford, New York.

Marlatt, G.A., 1992. Substance abuse: implications of a biopsychosocial model for prevention, treatment, and relapse prevention. In: Gabowski, J. (Ed.), Psychopharmacology: Basic Mechanisms and Applied Interventions. American Psychological Association, Washington, DC, pp. 131–162.

Marlatt, G.A., Baer, J.S., Donovan, D.M., Kivlahan, D.R., 1988. Addictive behaviors: etiology and treatment. Annual Review of Psychology 39, 223–252.

Newcomb, M., Earleywine, M., 1996. Intrapersonal contributors to drug use. American Behavioral Scientist 39, 823–837.

Shaffer, H.J., LaPlante, D.A., LaBrie, R.A., et al., 2004. Toward a syndrome model of addiction: multiple expressions, common etiology. Harvard Review of Psychiatry 12, 367–374.

Relevant Websites

http://www.cdc.gov – CDC
http://www.niaaa.nih.gov – NIAAA
http://www.nida.nih.gov – NIDA
http://www.nida.nih.gov/nidamed – NIDAMED
http://www.samhsa.gov – SAMHSA
http://en.wikipedia.org – Wikipedia

CHAPTER

7

Disease Model

Arthur W. Blume, Desiree M. Rudisill, Shai Hendricks, Natalia Santoya

Washington State University, Vancouver, WA, USA

OUTLINE

HISTORY AND OVERVIEW

The disease model views addiction as a physiological disease with specified symptoms and course. The disease model is a biomedical model that understands addiction as a physiological pathology that requires medical or pharmacological interventions. Because the disease model defines addiction as a sickness, proponents of the model use medical terminology to describe important addiction processes and treatment. The term treatment (rather than therapy, as an example) is used for the provision of health care services meant to intervene on addiction. Treatment modalities are defined medically (such as inpatient, outpatient, and day treatment). People who seek treatment are usually referred to as patients (rather than clients). Diagnostic definitions also reflect a disease model understanding of the nature of the problem (e.g. full and partial remission from the *Diagnostic and Statistical Manual* published by the American Psychiatric Association).

In the model, addiction is defined as a chronic disease for which there is currently no medical cure. The model also views addiction as a disease with symptom progression; not only is the condition chronic and incurable, but also the symptoms progressively worsen over time. Severity of the disease is often determined by the progression of various physiological symptoms, such as tolerance and withdrawal. Symptom progression is viewed by some proponents of the disease model as continuing in a latent fashion even when a patient is abstaining, so if the patient were to return to active addiction, it is likely that a progression of symptoms (when compared to when they had abstained) would be observed almost instantaneously upon return to the addictive behaviors even if the person had abstained for a significant amount of time.

In addition to these worsening symptoms over time, two physiological processes perpetuate the course of the disease: cravings and loss of control. Cravings as defined by the disease model are physiological responses often triggered by withdrawal processes that create a strong desire in the patient to seek out and use the substance. Loss of control is a disease process that contributes to an inability of patients to predict or control how much or how long they will engage in the addictive behavior (e.g. how much or how long they will use during a drinking or drug using event). Loss of control has also been referred to colloquially as powerlessness over a substance (first by members of Alcoholic Anonymous (AA)), and that powerlessness tends to cause unmanageability in everyday life (from the First Step of AA).

The disease model represents the first serious attempt to define addiction processes within the spirit of scientific inquiry. In the United States, Dr Benjamin Rush was the first to recognize the public health risks

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00007-3

associated with excessive alcohol consumption and to discuss alcohol abuse as a medical condition. He referred to the possibility of progression of alcohol symptoms in his model (1785), but his model suggested moral solutions rather than medical treatment when considering how temperance would be attained. In 1849, Magnus Huss, a Swedish physician, is credited with first referring to alcoholism as a disease that could result from heavy drinking of alcohol.

The disease model as it is generally understood and practice today was largely due to the work of American E. Morton Jellinek in the mid-twentieth century. Jellinek investigated the symptoms of alcoholism among a limited sample of chronic inebriates in an effort to determine common symptoms and course of the disease that could be used to plot progression and determine prognosis. From this investigation, Jellinek created a chart now referred to as the Jellinek curve that mapped out the progressive course of the disease of alcoholism symptom by symptom as determined from the interviews discussed above. The Jellinek curve (so named because of the U or V shaped curve of progression of symptoms on the chart) was also classified into progressive stages of severity in the disease based upon the progressively worsening symptoms, with the first stage referred to as prealcoholic, the second as prodromal, the third as crucial, and the fourth and last stage referred to as chronic alcoholism. The bottom of the curve (often colloquially referred to as rock bottom) leaves alcoholics with few options: imprisonment, insanity, death, or recovery (which then leads to the ascending side of the U or V shape as patients incur more recovery time and show symptom improvement).

Interestingly, Jellinek later revised his understanding of the disease of alcoholism in the 1960s, indicating that his research had suggested multiple presentations of alcoholism, rather than only one trajectory as first suggested in his curve. He subsequently described different typologies of alcoholism: alpha (psychological but not physical dependence), beta (chronic abuse that leads to health problems but may not include physical dependence), gamma (similar to the symptoms and course described in the curve), delta (similar to gamma but without loss of control), and epsilon (binge patterns). Jellinek considered gamma and delta alcoholism to operate under disease principles, but was less certain about whether the other typological categories did.

American Medical Association Defines Alcoholism as Disease

In 1954, the American Medical Association (AMA) publicly declared that alcoholism was a disease. This was a watershed event in the treatment of addictions and at the time was a radical, compassionate act that paved the way for third party reimbursement for the treatment of addictive disorders. Before this declaration, addiction was mostly viewed as a moral problem by the general public and government policies treated those with addiction accordingly. The declaration of the AMA not only increased the likelihood of effective treatment services, but also reduced some of the stigma in seeking help for alcohol-related problems. The declaration also put into motion subsequent legislation that allowed for federal funds to be used to support treatment and eventually provided funding streams for researchers investigating addiction. Although some critics have suggested that the AMA declaration was made more out of professional self-interests than compassion, the declaration created the necessary conditions for addressing addiction as a health problem rather than a moral problem in the United States.

The declaration also contributed to the rise of the treatment industry in the United States. With the onset of third party reimbursement for treatment, many hospitals and agencies saw the possibility of treatment as a profitable business as well as providing a community health service. Health care entities with exclusive specialization in addiction treatment delivery were constituted. Today, the treatment industry remains a large force in health care delivery in the United States with literally thousands of treatment centers generating revenues in the billions of dollars.

Most treatment facilities today practice what has been referred to as the Minnesota Model of treatment. The name is derived from the pioneering work of treatment centers in the state of Minnesota, such as Hazelden and Johnson, which as industry leaders developed a unique integration of services incorporating ideas and principles from both the disease model and the Alcoholics Anonymous/Narcotics Anonymous spiritual-based recovery model to determine the nature of therapeutic process of treatment (discussion of disease allied with work on the 12 steps and participation in recovery groups). It is worth noting that this was an unusual merger of conceptual models in what amounted to treating a biomedical disease with a spiritual recovery program. Since the disease model views addiction as incurable, proponents were forced to look outside of medical care for help in treating the problem. The logical source for finding help would be the existing 12-step recovery programs, which at the time of the rise of treatment centers would have been the only high-profile programs with a track record of success for some people. Today the Minnesota Model of treatment continues to be the most widely practiced in the United States, and large high-profile institutions such as Hazelden and Betty Ford tend to mostly operate under disease model principles today.

It is important to keep the timing of the AMA declaration in historical context. National prohibition had been the law of the land only 20 years earlier. Although some

of the impetus for prohibition legislation was centered on health concerns related to drinking alcohol, most of the effort was because of concerns about national morality and the belief that alcoholism was contributing to the moral degradation of the United States. The United States has a rather unique history when it comes to public policies related to addiction, including its somewhat lengthy history of experimenting with national prohibition. This unique history also contributed to the preeminence of the disease model as the theoretical framework for the treatment industry in the United States. The disease model arguably arose as a reaction to the generally uncompassionate and widely professed moral model beliefs that viewed addictive behaviors as a sign of a significant moral flaw (often thought of as sin) in a person and viewed behavior change as a matter of personal choice and willpower. The disease model challenged those beliefs by describing addiction as a disease that contributed to loss of control. Loss of control meant a person had no choice in the matter of addictive processes and was in fact sick rather than fundamentally flawed.

Other nations outside the United States did not necessarily experience the same history, and therefore did not have the need to challenge a moral model belief system. Some have referred to the disease model as uniquely American and its use beyond the United States is not widespread. The disease model continues to be predominantly practiced in most treatment centers in the United States, although other recovery models have made inroads as science has demonstrated the efficacy of their treatment techniques (cognitive behavior therapy in particular).

A Disease of the Brain

In more recent years, the belief that addiction is a disease has evolved somewhat to reflect the growing body of research in genetics and neuroscience. There are many that believe that addiction may be genetically predetermined (running in families) and that genetic research has demonstrated that addiction is a disease with hereditary risk (like heart disease for example). Neuroscientists and neurobiologists have determined physiological mechanisms of addiction and how addiction specifically impacts brain function. The new empirical findings have led many to conclude that addiction is a disease of the brain. The advancement in understanding the biological processes related to addiction over the last decade or two has been nothing short of miraculous, and it is increasingly clear that complex physiological processes are involved in the development and maintenance of addictive behaviors. It also seems clear that people who may be at the highest risk for developing addictions often have family histories for addiction, suggesting genetic and biological predispositions for risk of onset.

However, as many scientists have noted, there seems to be a large gulf between where the science of the field is today and what is being delivered clinically to patients in treatment centers. With regard to the disease model, most treatment centers continue to operate under disease model assumptions that in some cases have evolved little since first formulated 70 years ago. In addition, much of the new information being discovered by neuroscientists and neurobiologists is not being incorporated into treatment protocols. The result is that the evolving brain disease model being discussed among researchers is generally significantly different and more sophisticated than the traditional disease model philosophy being transmitted by treatment centers to their patients.

STRENGTHS OF THE DISEASE MODEL

The many strengths of the disease model have contributed to the evolution of addiction research and treatment. First, the rise of the disease model allowed compassionate health care to be provided to patients with addictive behaviors and improved accessibility to treatment services. Second, the disease model acknowledges the importance of biological processes at work in addiction. Third, related to the second point, the disease model has contributed to an explosion of biomedical research related to understanding and treating addictive behaviors. And fourth, use of the modified disease model by patients in recovery has improved the quality of the lives of many people struggling with addictions. Perhaps, the greatest legacy of the model is that it has provided hope for many people, especially during a period of history when none existed. The disease model was a compassionate alternative when first proposed as compared to the moral model. Many success stories in treatment and recovery have been documented among people who sought treatment services in facilities that operate under the assumption that addiction is a physical disease.

CRITIQUE OF MODEL

One important critique concerning the original conceptualization of the disease model involves the methods by which the model was developed. Critics of the model have pointed out that the small studies Jellinek used to determine the symptoms and course of the disease of alcoholism were populated by people who were essentially outliers with severe addiction symptoms. Studies involving outliers would not necessarily be applicable to all people engaging in addictive behaviors. Another criticism is that the symptoms and trajectory of the disease were qualitative interpolations of

what research participants reported retrospectively rather than deduced by empirical means with use of prospective data and rigorous experimental methods. Another critique of the methods has been that since the symptoms and course were assessed in study participants retrospectively, there may have been errors in what research participants recalled. Especially given the severity of addiction in the study participants, where memory problems would be of concern, participants would likely be poor historians with regard to what symptoms were experienced when. As mentioned, Jellinek moved away from the original model later in his career with his research about different typologies, suggesting that he too had lacked faith in it. However, the Jellinek curve continues to be used by many disease model proponents in clinical settings even though it was developed for use 70 years ago.

Another critique comes from those who point out that not everyone has benefitted from Minnesota Model treatment based on the principles that addiction is a disease that contributes to powerlessness and unmanageability. As an example, there have been estimates that of the millions of patients who have received treatment for addictive behaviors, only a small proportion have membership in the 12-step recovery groups after treatment as recommended by Minnesota Model centers. Some research has demonstrated that certain groups, such as ethnic minorities and women, do not associate with 12-step recovery programs in the numbers that would be expected, and 12-step fellowship groups tend to be more heavily populated by men and European Americans in general. There is also evidence that people who do not view spirituality or religion as important are less likely to seek out the 12-step recovery programs. In addition to these groups that have difficulties with the 12-step recovery culture or philosophy, there are also people who do not view their own addictive processes as chronic or progressive, and may not accept that they are powerless. People with these views often are uncomfortable with the traditional Minnesota Model treatment and unlikely to seek out 12-step recovery. Some of those who have not benefitted from traditional Minnesota Model treatment have found help in therapy that uses other models of recovery (cognitive-behavioral, for example).

Research concerning the success of treatment for those who do seek traditional Minnesota Model services has yielded interesting results. To be certain, treatment works, but not often in the way that would be predicted by those who are proponents of a chronic progressive disease. Researchers have found that abstinence rates for patients after treatment tend to be modest, and for a number of years those findings contributed to some concern that treatment may not be as effective as professionals would like. However, when treatment outcomes were examined by including data about reduction in addictive behaviors and improvement in quality of lives along with data concerning abstinence, the impact of treatment became clearer. It was discovered that many clients who graduate from traditional treatment programs that insist on abstinent goals were opting for reductions rather than abstinence and simultaneously reporting improvements in the quality of lives without abstinence. Treatment is indeed successful for many, if not most, of clients, but not in the manner predicted by the disease model. Critics of the disease model have concluded from these findings and other parallel research that patient's perspectives about their own addictive behaviors may be more important to consider rather than assuming the disease must be treated in a certain way in order for recovery to occur.

Relevant to these findings is new research on behaviors that have been typically labeled as denial. Research that contributed to development of Motivational Interviewing, an empirically supported intervention for addictive behaviors, revealed that what is often called denial/resistance in treatment was actually a function of interpersonal interactions between client and therapist. Researchers found that resistance and denial changed in a therapy session as a result of the type of strategies used by the therapist; increased confrontation led to increased coding of denial and resistance by behavioral observers, and denial/resistance would decrease when the therapist switched to a less confrontative style. Neuroscientists have investigated similarities between what was referred to as denial with neurocognitive impairment subsequent to addiction (memory problems, anosagnosia). Today, many researchers believe that what has been typically been referred to as denial, a term that has become arguably judgmental, would be better described as lack of trust in the interpersonal dynamic, lack of awareness of problems or the need to change, or perhaps indicative of a cognitive/perceptual problem.

Research on cravings has demonstrated how they can be triggered psychologically as well as physiologically. Scientists and clinicians alike had been baffled for years about why cravings would occur long after physiological withdrawal would be expected to have ended. It became increasingly clear that cravings also can be learned behaviors, classically conditioned and linked to specific cues in the environment that trigger the craving response. In addition, cravings also can be triggered by cognitions (thoughts, beliefs) such as positive outcome expectancies. The disease model has long conceived of cravings as the consequence of physiological processes, but substantial evidence now exists that cravings can occur independently of physiological processes and cravings can be effectively intervened upon by psychological means.

In addition, there is a substantial body of research that has demonstrated that loss of control may not

always occur among people with addictions. For example, scientists have determined that subgroups of people with addictive behaviors are able to mature out of the behaviors (young adults) as they age or are able to experience natural recovery without treatment (older adults). Some people who have met the criteria for substance dependence have been able to return to light or moderate engagement in the addictive behavior without demonstrating loss of control subsequently. Psychological studies, such as the one conducted by G. Alan Marlatt referred to as the balanced placebo drinking study, have demonstrated that people with alcohol dependence do not necessarily lose control of substance use when drinking alcohol, and that for some people loss of control may be more a function of beliefs about drinking circumstances than physiological responses. Different typologies of addiction may help to explain why some do not lose control or can regain control, but the empirical support for maturing out and natural recovery tends to dispute traditional assumptions inherent in the disease model.

Other research has demonstrated the potency of psychological and social/environmental factors on addictive processes in a way that would not necessarily be predicted for a biomedical condition. Psychological principles concerning reinforcement and cognitions have been found to be particularly effective in predicting not only subsequent addictive behaviors but also success in treatment. It is also known that interpersonal and environmental factors are associated with the risk for and predictive of the course of addictive behaviors. Like many other disorders, many researchers believe that addictive behaviors are thought to be best understood within the context of a biopsychosocial model that accounts not only for physiological processes such as genetics and the reward pathways of the brain, but also for psychological and social/environmental processes known to be associated with addiction. Belief in addiction as a brain disease must adequately account for the psychological functions inherent in the human brain as well as physiological processes.

The disease model tends to heavily focus on a biomedical approach to understanding and treating addictive behaviors in spite of the significant evidence that addictions are more appropriately understood as biopsychosocial processes. In defense of the biopsychosocial model, many professionals have noted how patients in disease model treatment have often been derailed in their recovery by psychosocial stressors, and in fact many Minnesota Model treatment centers have incorporated psychosocial interventions (often cognitive-behavioral such as relapse prevention) to accompany biomedical and 12-step interventions. Critics have suggested that the evolution of Minnesota Model treatment over time to include psychosocial interventions indicates that biomedical and 12-step model provided insufficient means alone for addressing many of the psychosocial factors that impact addiction, and this is an evidence of the importance to view addiction from a biopsychosocial perspective.

One final note is that the disease model represents a Western view of medicine that does not necessarily match well to other world perspectives on medicine and illness. The disease model has not been widely accepted or practiced in areas of the world where other models of medicine and illness exist. The disease model also has not been widely accepted among immigrants to the United States from areas of the world who have different perspectives of medicine and illness than those posited in the disease model. The difference in beliefs regarding medicine and illness has in some instances created a mismatch between client worldviews in treatment and the model of the treatment provider providing care. In other instances, people with different views of medicine and illness will simply avoid the use of traditional treatment services, usually seeking out cultural relevant methods instead.

POSITIVE CONTRIBUTIONS OF THE MODEL

In spite of criticism, the original disease model represented a qualitative improvement over existing models when first described by Jellinek and others. Addiction was often viewed as a moral problem and therefore compassionate care was lacking. People with addiction were often sent to asylums or spent time in prison. Only after the advent of the disease model was there interest and means to treat people for their addictions in a health care setting. Addiction became a diagnosis rather than a moral problem, and research was stimulated to find treatments and cures. The disease model was radical when proposed: It represented a compassionate reconceptualization of addictive processes. Years of subsequent research have revealed that addictive processes are much more complex than initially realized, but without the first conceptualization of addiction as disease, that empirical inquiry likely would not have occurred.

The introduction of the disease model and subsequent declaration by the AMA in 1954 paved the way for third party reimbursement for treatment. The disease model not only made treatment possible, but also increased its accessibility to those in need. Without third party reimbursement for addictive disorders, the number of treatment centers in the United States would likely be greatly reduced in numbers and potentially even more expensive than found today.

SUMMARY AND FUTURE DIRECTIONS

The disease model was the first modern conceptualization of addictive behaviors. Because it was the first empirically derived model, its legacy is immense. Without the development and acceptance of the disease model, both clinical practice and research would look much different today, especially in the United States. As discussed, the model as traditionally conceived has its flaws and limitations. However, the disease model has stimulated many promising areas of inquiry into addictive behaviors, including in medicine, neurobiology, neuroscience, pharmacology, and the social sciences. The future of the disease model is difficult to determine. On the one hand, significant biomedical research gains have been made that support the strength of association between the brain and addictive processes. On the other hand, many supporters of the disease model have focused so intently on the biological processes of addiction that they have lost sight of the importance of psychological and social/environmental processes in understanding and treating addiction. It is difficult to conceive of a model effectively defining and treating addictive behaviors in the future that ignores its psychological, social, and environmental factors. It seems likely that treatment centers will continue to use pragmatic methods to improve care by cobbling together treatment techniques that are derived from many different models for understanding addiction, perhaps suggesting the disease model will evolve with the introduction of these new treatment ideas.

The gulf between how researchers define addiction as disease and how practitioners define addiction as disease is immense and presents the greatest challenge. At all but a very few select treatment centers, the disease that is discussed is derived from information largely out of date. In addition, a substantial amount of what is practiced in many traditional Minnesota Model treatment centers today has not been empirically validated as effective or demonstrated to be best practice. The greatest challenge for disease model proponents in the future will be to bridge the gap between science and practice in a way that greatly accelerates the benefits that patients receive as the result of new scientific discoveries.

SEE ALSO

The Biopsychosocial Model of Addiction, Cognitive Factors in Addictive Processes, Denial and Lack of Awareness in Substance Dependence: Insights from the Neuropsychology of Addiction, Contextual Factors in Addiction, Historical Understandings of Addiction, Maturing Out, Natural Recovery

List of Abbreviations

AA Alcoholics Anonymous
AMA American Medical Association
NA Narcotics Anonymous

Glossary

Alcoholism colloquial term often used with disease model to describe alcohol dependence.
Controlled drinking the ability of some people to return to some use of alcohol use without loss of control after being diagnosed with alcohol dependence.
Biomedical model belief that disorders like addictions have their root causes in biological mechanisms and that the best treatment for such disorders is medical.
Compulsion an irresistible physiological impulse to engage in the addictive behavior and contributes to loss of control
Denial originally a defense mechanism in Freud's model, it has been widely used to describe when patients refuse to acknowledge powerlessness over addiction over take action to change their behavior.
Disease model model that posits that substance abuse operates under biomedical principles and is a chronic and progressive disease with prescribed symptoms and predictable course.
Genetic predisposition belief that a person will be at increased risk for addiction because of genetic loading.
Natural recovery many people are able to spontaneously recover from addictions without treatment or therapy.
Powerlessness (loss of control) concept originally derived from the first of 12 steps of Alcoholics Anonymous; the belief that addictions contribute to complete loss of control of the behavior that leads to unmanageable lives.

Relevant Websites

http://www.aa.org/ – Alcoholics Anonymous.
http://www.bettyfordcenter.org/ – Betty Ford Center.
http://www.hazelden.org/– Hazelden Treatment Center.
http://www.tgorski.com/ – Terence Gorski Web Site.

Further Reading

Alcoholics Anonymous, 2001. Alcoholics Anonymous, fourth ed. Alcoholics Anonymous World Services, New York.
Jellinek, E.M., 1942. Alcohol Addiction and Chronic Alcoholism. Yale Press, New Haven, CT.
Jellinek, E.M., 1960. The Disease Concept of Alcoholism. Hillhouse Press, New Haven, CT.
Marlatt, G.A., Demming, B., Reid, J.B., 1973. Loss of control drinking in alcoholics: an experimental analogue. Journal of Abnormal Psychology 81, 233–241.
Midanik, L.T., 2006. Biomedicalization of Alcohol Studies: Ideological Shifts and Institutional Challenges. Aldine Transaction, New Brunswick, NJ.
Miller, W.R., 1993. Alcoholism: toward a better disease model. Psychology of Addictive Behaviors 7, 129–136.
Peele, S., 1999. Diseasing of America: How We Allowed Recovery Zealots and the Treatment Industry to Convince Us We are Out of Control. Jossey-Bass, New York.
Peele, S., Brodsky, A., 1992. The truth about addiction and recovery. Fireside, New York.
Valliant, G.E., 1995. The Natural History of Alcoholism Revisited. Harvard University Press, Cambridge, MA.

CHAPTER

8

Denial and Lack of Awareness in Substance Dependence: Insights from the Neuropsychology of Addiction

A. Verdejo-García[*], *M.J. Fernandez-Serrano*[$], *J. Tirapu-Ustarroz*[**]

[*]Institute of Neuroscience F. Olóriz, Universidad de Granada, Granada, Spain, [$]Department of Psychology, Universidad de Jaén, Jaén, Spain, [**]Clínica Ubarmin, Egüés-Navarra, Spain

OUTLINE

INTRODUCTION

Substance use and dependence is characterized by a difficulty to control consumption and a tendency to persistence despite its increasing adverse consequences (DSM-IV, CIE-10). Current neuropsychological models conceive addiction as a neuroadaptive process leading to significant alterations in the frontal-striatal circuits, including several sections of the prefrontal cortex (orbitofrontal, dorsolateral, and medial), limbic and paralimbic regions (amygdala, hippocampus, insula), and basal ganglia. Alterations in these circuits affect the activity of the motivational, emotional, and executive control systems. Dysfunction of motivational systems is made evident by the overestimation of the reinforcing, adaptive value of drug-related stimuli and the underestimation of the value of other natural reinforcers, fundamental to our survival. This motivational bias promotes the development of impulsive behaviors, which are fostered by the expectation of immediate reinforcement (drug use–related), whereas little consideration is given to its potential, long-term consequences. Continuous exposure to drug use turns these impulsive behaviors into habits, that is, impulsivity gives way to compulsion, thus consolidating motor programs that go off automatically in the presence or anticipation of drug-related signals. Repeated consumption also provokes important changes in the interoceptive and emotional systems. Recent studies highlight the role that detection of body signals related to drug-taking experiences plays in the generation of emotional states leading to drug-seeking behavior, such as craving. In addition, a persistent sensitization of the stress systems takes place and the drug-seeking behavior rockets in an attempt to rebalance the body's hedonic tone. Alongside this adaptation of the motivational and affective systems, addiction is associated with a deterioration of the executive control mechanisms, which depend on the functioning of the prefrontal cortex. The dysfunction of executive mechanisms leads to alterations of processes which are considered of a more cognitive nature, such as the ability to update and monitor multiple sources of information (cognitive, affective, and motivational) aiming to design plans and reach

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00008-5

goals. In addition, it hinders the ability to inhibit and effectively switch response patterns according to the changing demands in the environment, as well as the capability for making adaptive decisions that guarantee both our survival and our survival quality.

The notion of lack of awareness of addiction and its consequences is inherent in a number of the neuropsychological dysfunctions reviewed above. Neuroadaptations in interoceptive, motivational, and affective systems may contribute to a persisting attentional bias on drug-related needs that prevents considering other homeostatic signals. The results of recent neuroimaging studies show that dysfunction of the insula (a key region for processing and integrating interoceptive signals) may underlie the difficulties of addicted individuals to gain access to the emotional signals that are necessary to realize the implications of the disorder and develop alternative behaviors. In addition, the consolidation of an impulsive response pattern implies a lack of consideration of the input associated to delayed information, which would include most of the negative consequences of addiction. In this sense, we know that the ability to imagine future emotional events is related to the functioning of the medial orbitofrontal cortex, one of the most affected regions by addiction. On the other hand, once the addictive process is established, many processes oriented to drug seeking and taking (e.g. programming of motor habits) may operate in a sophisticated manner without full conscious supervision or control. Finally, dysfunction of executive mechanisms affects mainly the individual's ability to maintain the correspondence between intention and action and to compile all necessary information to become aware of his/her deficits and their repercussions on the family milieu and social environment. To sum up, to be aware of the symptoms and repercussions of any disorder, we need to have access to all relevant information (internal and external) and be able to properly compile and monitor this information to understand its implications and use it to establish goals and implement change-oriented behaviors. In addiction, however, we find that many processes underlying these capacities are significantly deteriorated. In these cases we may talk about lack of awareness – but not denial as a meta-cognitive deficit resulting from deterioration of information processing and integration at different levels. This deterioration is linked to the dysfunction of frontal-striatal circuits involved in addiction.

The concept of denial implies an implicit knowledge that can be accessed depending on the state of the organism in a particular situation, as well as an elaboration process aimed to minimize or suppress this input. Parallel neuroadaptive processes typical of addiction can also help understand this phenomenon. In their extreme form, denial processes could be compared with the symptoms of confabulation observed in neurological patients. Symptoms of confabulation are common in patients with damage in the medial orbitofrontal cortex, one of the key brain regions both in self-referential processes and in addiction. There is evidence that confabulations in these patients are related to (not necessarily conscious) efforts to maintain a positive self-concept or a self-referential sense of coherence, thus building a valid argumentation to maintain a status quo that will not put their own identity at risk. These confabulation mechanisms oriented to maintain a status quo could help explain the phenomenon of denial in addicted individuals. On the other hand, partial access to input regarding disorder implications may exacerbate the reactivity of the stress systems, leading to intense discomfort and undermining the individual's self-efficacy expectations on his/her capacity to cope with the problem. This unbearable rise in stress levels makes the individual react by expelling this information from his/her conflict resolution system. These psychological processes would depend on the dialog between the brain regions responsible for the motivation and conflict (anterior cingulate cortex) and the neuroendocrine systems regulating the response to stress (hypothalamic–pituitary–adrenal axis), which are persistently altered in the addictive processes. Other authors have also highlighted the role of the cross-talk between the cognitive styles of both brain hemispheres (right hemisphere, specializing in processing novelty vs. left hemisphere, specializing in contextualizing information according to preestablished patterns) as a substrate of these dissociations. Finally, a milder form of denial may result from the predominance of motivational resources driving drug seeking and taking on those resources in charge of motivation to change. In these cases patients would be partially aware of their disorder and the need for change, but they would tend to minimize or defer it. Classical models on stages of change claim that denial would be a process inherent to the precontemplation stage. In this stage the individual considers the possibility of changing his/her drug use behavior but may deny the problem in different manners, including resignation due to lack of self-efficacy necessary to produce change, or deferral of the change. In agreement with this notion, recent studies indicate that drug users present decision styles characterized by procrastination.

In sum, we understand lack of awareness and denial are two separable neuropsychological processes. However, we believe that they both have underlying neuropsychological (motivational, emotional, and meta-cognitive) alterations related to addiction-related neuroadaptations in frontal-striatal circuits. In the following sections we shall: (1) define a neuroscientific conceptual framework to understand the phenomena

of lack of awareness, denial, and lie in addiction; (2) describe the relationship between neuropsychological processes and symptoms of anosognosia and denial; and (3) review the empirical studies having dealt with this phenomenon from different approaches over the last decade.

AWARENESS, LIE, AND DENIAL IN ADDICTION: A NEUROSCIENTIFIC APPROACH

Self-Awareness

Being conscious and the conscious experience may be explained by models based on the more proximal brain function, related to general neuronal action patterns. However, we all agree that awareness is more than that. When we say that someone is not aware of something, we are referring not only to tacit knowledge, but also to other relevant aspects, such as the implications of that knowledge. In the case of addiction, when we talk about lack of awareness, we refer not only to aspects such as insight of the cognitive deficits associated with drug use, but also to awareness itself about suffering this disorder. The latter would be more closely linked to the concepts of self-awareness and self-concept, which use defense or self-protection mechanisms to avoid the emotional consequences of accepting that one suffers a disorder. Self-awareness has been defined as a human tribute that allows not only awareness of the self but also awareness of one's position in his/her social environment. In a hierarchical organization of mental functions, the ability of self-awareness would be in the vertex of the pyramid, since its function is to control one's own mental activity, representing current experiences in relation to previous ones, using acquired knowledge to solve new situations, or guiding decision making for the future. According to Prigatano, self-awareness is the ability to perceive oneself in relatively objective terms, maintaining a sense of subjectivity. This aspect of self-awareness implies a cognitive process, as well as an emotional state, and the integration of both.

The critical neural system for self-awareness is located in the prefrontal cortices, whose neuroanatomic position favors this purpose because (1) they receive signals stemming from all sensory regions where conscious experiences are formed (including images making part of our thoughts); (2) they receive signals from somatosensory cortices representing past and current body states; (3) they receive signals from bioregulatory sectors of the brain, among which neurotransmitters of the brainstem and basal prosencephalon, as well as the amygdala, the anterior cingulate, and the hypothalamus; and (4) they categorize the situations that affect the body, that is, they classify any contingencies in our life experience. Therefore, the prefrontal cortex is often a convergence zone that works as a deposit of dispositional representations for properly categorized, unique contingencies of our life experience. In this sense, it seems evident that, to exist, self-awareness needs information from signals coming from both outside of the self and our thoughts (which are also images). Also, it needs to have access to previously categorized information (the categorization criteria are probably based on the emotional value of the experiences). At this point, we create representations that become unique dispositions, as they are based on our life experience, which is also unique.

From this point of view, and according to Stuss and colleagues, the alterations of awareness related to lesions in the prefrontal cortex have some peculiarities that should be taken into account: (1) these alterations are connected to the self and become evident generally in behavioral rather than cognitive functioning; (2) these alterations may occur irrespective of any cognitive or sensory deficits, even with an intact IQ; (3) executive functions (i.e. higher order abilities involved in the generation and control of goal-directed behavior) are important, since they are more specifically related to prefrontal systems and their damage may be linked with the alteration in awareness of deficits in behavioral functioning; (4) there may be a fractioning of awareness in relation to specific connections between the prefrontal cortex and other specific regions; (5) self-awareness is more than mere knowledge, it is the ability to reflect on the implications of this knowledge; and (6) self-awareness does not only refer to the past and the present, but it also projects to the future. In this sense, we would depict awareness not only as knowledge, since knowledge does not have implications if devoid of emotional valence. In the case of addiction, structural and functional alterations in the prefrontal cortex have been described by several neuroimaging studies in abusers of a number of drugs. Furthermore, the main corollaries of prefrontal cortex–related awareness alterations largely overlap with the neuropsychological alterations described in addicted individuals. Specifically, addicted individuals: (1) show a wide range of behavioral alterations associated with the functioning of the circuits connecting the prefrontal cortex with subcortical regions and basal ganglia (including symptoms of apathy, disinhibition, and behavioral disorganization); (2) show important alterations in executive functions, even in the presence of normal, general cognitive functions and IQ; (3) have problems to think about the consequences that these and other drug use-related alterations bring about in their daily functioning and their family milieu and social environment; and (4) show difficulties

in associating cognitive scenarios with appropriate emotional states, especially when it comes to anticipate emotional consequences associated with future events.

This more neuropsychological approach claims that there are different forms of deteriorated awareness, depending on the brain systems suffering the damage or dysfunction. In Fig. 8.1 we present a schematic representation of the different brain systems linked to awareness-related deficits. In this line, neuroscientific models have proposed in a generic manner that bilateral lesions of the frontal cortex are specifically associated with the lack of awareness on social behavior and executive functions, and with self-awareness as a whole. Clinical experience shows that there is a particular type of awareness for each knowledge module, which is fed by updated information from that particular module, whose goal is to guarantee the best adjustment of the individual. In this sense it seems that this awareness for each specific area is affected when the brain damage affects the function and its awareness locus. We must take into account that addiction-related brain alterations may interfere with the neural substrate of self-awareness. Therefore, if a prejudicial drug use pattern produces a deficit in the cognitive, emotional, and behavioral sphere that the subject is not aware of, this lesion affects both the locus of the relative processes and the locus of their awareness. In other words, it would affect the cognitive function, the emotional sphere or its behavioral patterns, as well as the metacognition of the cognitive, emotional, or behavioral function. We could state, in a graphic manner, that the individual does not know that he/she does not know and this first aspect would make reference to concepts such as change of personality, suffering provoked to others, empathy, etc. In addition, awareness for a specific knowledge module may be affected by the degradation

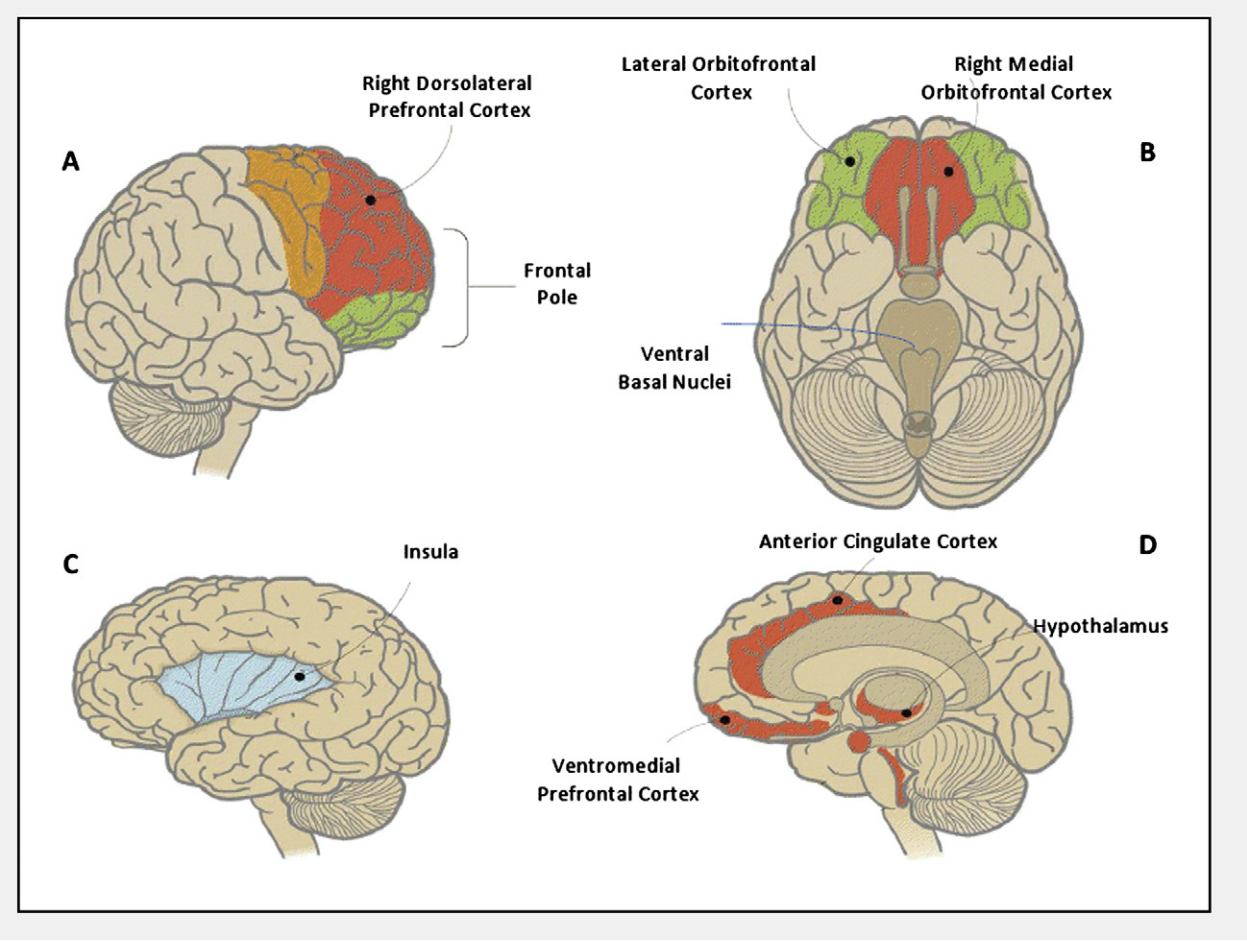


FIGURE 8.1 Schematic representation of the main brain regions supporting awareness and denial-related deficits in substance-dependent individuals. 1. Right dorsolateral prefrontal cortex (Panel A) and right medial orbitofrontal cortex (Panel B), along with insula (Panel C), are involved in deficits of lack of awareness. 2. Interhemispheric right-left cross-talk or connections between anterior cingulate cortex and stress systems stemming from hypothalamus mediate processes of denial (Panel D). 3. Medial/lateral orbitofrontal cortices and frontal pole are key structures for the emergence of confabulations (Panels A and B). 4. Dorsolateral prefrontal cortex and dorsal anterior cingulate control are necessary systems for lie (Panels A and D).

of feedback of its response or output or because the internal representation of the cognitive processes is also affected as a result of the brain impairment. For example, drug use may affect attentional mechanisms, which would lead to a degradation of the information sent to the brain. Thus, the subject cannot be aware due to the affectation of input mechanisms (Wernicke's aphasia may be the clearest example in neurological patients). In other words, we find that drug use leads to the affectation of cognitive, emotional, and behavioral functions, as well as the affectation of cognitive, emotional, and behavioral mechanisms implied in self-knowledge and self-awareness.

Denial

Some studies have proposed the existence of different kinds of alteration of awareness of one deficit according to the three basic levels of information processing: information collection, neuropsychological, and emotional. Thus, the lack of awareness itself would be the result of the patient not receiving information or this information being degraded. Also, it is possible that the patient does not understand the meaning of the information (e.g. it has been proved that alcoholic people have difficulties in processing faces expressing some basic emotions, such as sadness, relevant to understand damage inflicted to others). Regarding the level of implication of this information, the subject cannot take the self as an object or understand the implication of the deficit. Additionally, the arousal level for awareness may also be poor (a phenomenon frequently observed in benzodiazepine or opiate addicts). In this case, the deficit occurs at the information processing and access level. However, there are other phenomena related to the lack of awareness of different nature and characteristics. The phenomenon of minimization is characterized by the fact that the patient cannot understand or extract his/her own consequences and/or implications from the information. On an emotional level, the patient knows but cannot stand the effect of the information, thus reducing it to tolerance levels. Finally, in the case of denial, the patient has an implicit knowledge about the problem, but he/she cannot believe the information, which is too stressful, and expels it from his/her aware experience (*see* Table 8.1). On the other hand, studies have shown that denial observed in some chronic alcohol abusers does not necessarily need to be explained as a consequence of a maladaptive defense of the self, but as a manifestation of the neurotoxic effects of alcohol.

Prigatano has investigated the alterations in awareness of deficits in patients suffering brain damage, based on Weinstein's pioneer studies in the 1950s. Weinstein reckoned that the way that patients adapt and symbolically represent (a term with dynamic connotations) their deficits is determined by a series of factors, such as: (1) the type, severity, and location of brain damage; (2) the nature of the incapability; (3) the meaning that this deficit may have in terms of premorbid values and

TABLE 8.1 Nature of the Consciousness Problem Considering Three Possible Levels of Information Processing

Nature of the consciousness problem	Level 1 – Information	Level 2 – Implications (Neuropsychological)	Level 3 – Integration (Emotional)
Lack of awareness	- The individual cannot access the information. - The individual has not enough cognitive resources to understand the meaning of the information. - Main symptom is anosognosia.	- The individual cannot take the "self" as an object. - The individual cannot understand the information. - The individual cannot retain or remember the information. - Insufficient arousal for full conscious experience.	- No strong emotional implications.
Minimization	- No alterations at this level.	- The individual cannot extract from the information its consequences and implications.	- The individual knows but cannot tolerate the effect of the information, thus reducing it to tolerance levels.
Denial	- No alterations at this level.	- No major neuropsychological alterations although sensitization of stress systems and cross-talk between conflict control and stress systems or right vs. left hemispheres cognitive styles may contribute.	- The individual cannot believe the information, it's too stressing and therefore it is expelled from conscious experience.

experiences; and (4) the context in which the behavior is elicited and observed. According to Weinstein, denial must be understood as a loss of insight and is related to confabulation, lie, or symbolic disorientation. Also, Weinstein exposed the importance of premorbid personality factors to understand the mechanisms of denial, as well as the presence of implicit knowledge. In this sense, the mechanisms of denial are adaptive as long as they represent and explain the patient's incapability, giving it a sense of reality. In this line, Prigatano claims that denial of a deficit is due to several factors affecting superior levels of brain integration. In this sense, the cognitive capability of patients with brain dysfunction may be affected in evaluating the feedback about their functional limitations received from the environment. In addition, and despite receiving information about their own problems and limitations from the environmental feedback, they persist in maintaining a somewhat indifferent attitude toward this information. In fact, recent studies seem to prove that the insula plays an important role in the awareness of affective states, perhaps because the information on these affective states is produced by the brain mapping of associated body states. These connections may contribute to explain why in addiction profoundly altered body states and behavioral problems do not lead to full recognition of the disorder (*see* Fig. 8.1).

About Lie

Lie is a constant in approaching the phenomenon of drug addictions, contaminating very often the therapeutic relationship with patients. Lie does not refer to the consumption of toxics alone; deception goes beyond that and has to do with the perception of the problem. In fact, it is surprising that a patient denies having problems with his/her partner, children, or at work due to alcohol use (alteration of central awareness), and even more that he/she can create an unreal autobiographic story, where there seems to be no problem whatsoever (alteration of extended awareness). In therapeutic intervention, it is frequent to find patients that deceive or lie. In fact, distrust in patients' claims is the ultimate cause of taking urine samples for analysis.

The first question is: why do patients lie? The response is simple: what do they win by telling the truth? Cognitive psychology suggests that a basic stone for the treatment of drug addicts is anticipation of the consequences of their behavior, so that they will behave accordingly. This leads to a second question: to what point do they need to anticipate? The variable time is essential in brain functioning: if the anticipated consequences are negative in the short term, behavior inhibits the truth. It is a brain functioning mechanism: my brain has an image of what I am and how I am; a self-protecting image that compares data coming from experience to expel them if they contradict the image of myself, thus maintaining my *status quo.* There are patients that obviously know that they are lying, but they do it to protect themselves from the consequences of truth. When a subject generates mental images of the consequences of being honest, these images generate a negative emotion that inhibits the behavior, thus lie having an adaptive preservation value. Actually, it can be claimed that awareness is a human tribute that allows realizing the own reality and position in the social environment. Lie can, therefore, maintain the concept of reality and social status. In this sense, it can be claimed that the lack of awareness of a problem has the invented truth as a basic pillar to convert subjectivity into objectivity, so that this subjective, unique, private interpretation of my reality will be perceived by others in the way I perceive it. This will allow my maintaining the situation and more importantly the image I have of myself.

Trivers has suggested an ingenious explanation of the evolution of self-deception: in everyday life there are many situations where we need to lie. In this regard, researchers like Ekman have proved that liars always give themselves away with a somewhat feigned smile, an expression of tension, or a false tone of voice that can be detected by others. This is because the limbic system (involuntary and prone to tell the truth) controls the spontaneous expressions, whereas the facial expressions we make when we lie are controlled by the cortex (not only responsible for the voluntary control, but also the place where lies are invented). For Trivers, this problem has a solution: to effectively lie to another person, it is necessary to lie to oneself first. If we believe that the thing we are saying is true, our expressions will be real, with no trace of deception. However, this statement incurs into an internal contradiction, since it contradicts the proposal of self-deception, which implies that one can have access to the truth at any time; otherwise, self-deception would no longer be adaptive.

Ramachandran proposes a way to avoid this problem and states that a belief is not necessarily unitary: it is possible that self-deception is located in the left hemisphere, whereas the right hemisphere continues to know the truth. For this author, the key for self-deception lies in the working division between both brain hemispheres and our need to create a feeling of coherence and continuity in our lives. It is well known that the brain consists of two symmetrical halves, each of which specializes in different mental capabilities; the most noticeable brain asymmetry is related to language. Alongside these known function divisions, Ramachandran suggests the existence of an even more essential difference between the cognitive styles of both hemispheres, which can help explain the modalities of denial and lack of awareness. In this regard, Gazzaniga's

research on split brain is most revealing. This author has studied a series of patients with a disconnection between both hemispheres, which allows us to know the information stored by each of them separately. In Case P.S., one of the paradigmatic cases reported, the individual's left hemisphere is asked what he wants to do after finishing school, the answer being: I want to be a draftsman. However, when the right hemisphere is asked the same question, the answer was: automobile racer. This case, like others in the literature, proves that both brain hemispheres live different realities and suggests that everyone can have a mute brain inside the head with a reality and a perception of oneself very different to what we daily believe we are.

In any moment of life, the brain can feel overwhelmed by a continuous information cascade that must be integrated into a coherent perspective of the own image and the image that others expect from oneself. To generate coherent actions or to maintain a given *status quo*, the brain must have a type of mechanism that allows filtering and ordering this information in a stable, internally coherent scheme. This is what the left hemisphere is in charge of: integrating the information into the previous image of self. What happens when a piece of information on the own behavior does not fit into the script? The left hemisphere disregards this information completely or distorts it so that it fits into the preexisting frame to maintain the stability. These everyday defense mechanisms, far from being adaptive defects, prevents the brain from being doomed to incoherence and lack of direction due to the multiple combinatory possibilities of the scripts that can be written with the material collected by our experience. The problem is that one lies to oneself and to the others, but this is an efficient and affordable price compared with the coherence and stability that the whole system acquires.

The adaption strategies used by both hemispheres are basically different (*see* Fig. 8.1). The task of the left hemisphere is to create a model, a belief system, and to fit every new experience into this system. When it finds information that goes against this belief system, it turns to denial: it represses and invents a story that allows maintaining the *status quo*. However, the right hemisphere's strategy involves acting by calling this *status quo* into question, and searching for global inconsistencies. When the anomalous information reaches a certain threshold, the right hemisphere undertakes a global review of the model. In other words, the right hemisphere imposes a change of paradigm. This threshold is specific for each individual and depends on different aspects such as the features of personality or the type of experience. This would partly explain why after a stay in hospital for organic problems some individuals are able to accept their *status* as alcoholics, whereas others deny it at all costs, or why some individuals accept a relapse after the first time, while others need several relapses before they finally admit it.

Table 8.2 depicts a conceptual model of the psychological, psychopathological, neuropsychological, and relational implications of the three concepts presented

TABLE 8.2 Distinctive Features of the Different Forms of Reality Distortion Typically Found in Addicted Individuals

Features	Lie	Confabulation	Denial	Lack of awareness
Psychopathology	No	Yes	Yes/No	Yes
Intentionality	Yes	Yes/No	Yes	No
Consciousness	Yes	Yes/No	No	No
Type of knowledge	Explicit	Implicit	Implicit	Does not exist
Occurs within healthy individuals	Yes	No	Yes	No
Function	Deception	To make sense of a personal narrative	Self-deception	Don't have
Premeditation	Yes	No	No	No
Access to the truth	Yes	No	Yes	No
Plausibility	Yes	Yes/No	Yes	Yes
Receptor	The other	Oneself/the other	Oneself/the other	Oneself/the other
Neuropsychological impairment	No	Yes	No	Yes
Knowledge	Knows that he/she knows	Don't know he/she doesn't know	Don't know he/she knows	Don't know he/she doesn't know

in this section (lack of awareness, confabulation, denial, and lie).

SELF-AWARENESS, DENIAL, AND COGNITIVE DEFICITS

We have seen that neuroscientific evidence links the symptoms of lack of awareness and denial with cognitive and emotional alterations derived from a dysfunction of the frontal-striatal circuits. In this section we shall review the neuropsychological studies that have addressed this issue in a direct manner. These studies can be classified according to three methodological approaches: (1) studies on the connection between measurement of denial level in addicted individuals and their performance in neuropsychological performance tasks; (2) studies on the concordance degree between the results of objective performance measurement and the individual's insight of the own performance; and (3) studies on the concordance degree between the subjective information provided by addicted individuals and that provided by relevant informants regarding the problems derived from their addiction.

Regarding the first approach, Rinn et al. studied the connection between clinical estimations of denial levels in a group of alcoholic individuals under treatment and their neuropsychological performance on memory and executive function tests. At the beginning of the treatment they identified specific objectives related to denial symptoms that alcoholic subjects had to achieve during the test. To obtain a quantitative measure of the individual's denial level, they calculated the proportion of treatment objectives specifically for symptoms of denial that had not been completed. This measure was correlated with the performance indices from neuropsychological tests. The results showed that the degree of denial in addicted individuals significantly correlated to the impairment of the processing speed, memory, and executive function processes.

In the light of these results, the challenge lies in finding the nature of the relationship between these cognitive functions and the mechanisms of denial. Regarding processing speed and denial, these data could show that the speed to understand information from outside is affected in addicted individuals. On a psychological level, this impairment produces a degradation of the input arriving at the brain. On an anatomical level, it reduces the connectivity between different regions (processing speed deficits have been associated to alterations in the white matter). Both levels are essential to create a coherent, unified global image of oneself and the surrounding world. As for the relationship between memory and denial, we must understand that self-awareness is the capability of being conscious of a wide range of entities and facts, which generates a sense of individual perspective and space–time continuity. Therefore, it can be stated that we generate pulses of conscious experience for a goal, while generating an accompanying set of reactive autobiographical memories at the same time. Without these memories, we would not have a feeling of past or future, nor would a historic continuity exist (self-awareness). In other words, without conscious experience and memory, there is no self-awareness. At this level, we can locate confabulation phenomena oriented to reconstruct a sense of autobiographical coherence in addicted individuals. In fact, significant confabulation error rates have been observed in verbal memory tasks in psychostimulant users. Regarding the relationship between executive functions and denial, we consider that cold executive function processes can be more closely linked with the cognitive aspects implied in the insight, such as the updating of information about our own behavior or the cognitive flexibility that allows us change our rules and response patterns according to the feedback received. Therefore, they would be more closely related to the concept of knowledge of what is happening. However, there is a cognitive-emotional component necessary for the arising of insight, so that the subject can knock down his/her defense mechanisms (by saying defense, we do not adopt the dynamic vision, as we consider that each behavior has a brain correlate). In this sense, studies taking the Iowa Gambling Task as a paradigm highlight that addicted individuals have their decision-making processes affected, thus being guided by the immediate reinforcement and ignoring the long-term consequences of drug use. Furthermore, we know that this process is the result of the juxtaposition of cognitive processes and emotions, which takes place in the ventromedial prefrontal cortex, one of the regions that drug addiction most negatively effects. In this line, it is interesting that empathy of subjects with ventromedial affectation seems to be affected, which includes awareness of the potential damage inflicted to others. In addition, recent studies have proved the relationship between the activation of the ventromedial sector and self-knowledge. Both research lines offer a link between hot executive functions and self-awareness and insight.

The second approach consisted on studying the correspondence between measures of objective performance and the individual's insight on his/her own performance. Goldstein et al. have used a monetary effort attentional task (i.e. participants have to rapidly respond to get different amounts of money) in two groups: individuals with cocaine use disorders and healthy controls. The authors correlated participants' performance in the task with subjective measures of motivation toward the

monetary stake (a measure of state motivation). The results of these studies showed that, unlike healthy individuals, there was no correspondence between the subjective information provided by cocaine users about their level of motivation toward the monetary stakes and their actual performance in the task. Following a similar approach, several studies have shown that there is a certain level of dissociation between intention and action in addicted individuals. Two consecutive studies by Moeller et al. used a probabilistic learning task where participants (cocaine abusers and healthy individuals) had to choose between images with a pleasant, unpleasant, neutral or cocaine-related content. Additionally, subjective reports on the participants' preferences in these choices were collected. The results of both studies showed that, contrary to what happened in healthy individuals, cocaine abusers showed significant inconsistency levels between the images chosen and their subjective preference reports. One of this study also showed that this inconsistency was especially obvious in current cocaine users, as opposed to abstinent users. Noticeably, within the abstinent consumer group, the highest inconsistency levels were observed in individuals with higher cocaine consumption in the month before the onset of treatment. Similarly, Verdejo-García et al. used a self-regulation task where participants (cocaine-dependent and healthy individuals) had to discover on the go an implicit strategy that would allow them to improve their performance. Seventy percent of the consumers managed to correctly identify this strategy when they were debriefed about it. However, after analyzing their performance scores, these were significantly inferior to those of healthy individuals (i.e. despite having identified the optimal strategy, they had failed to implement it to optimize their performance).

Finally, a last approach is based on the analysis of the concordance degree between the subjective information provided by addicted individuals and the one provided by relevant informants to them (e.g. close relatives or friends) regarding addiction-derived problems. Following this approach, the study by Verdejo-García et al. examined the degree of awareness of substance-dependent individuals about their potential behavioral problems linked to alterations of frontal-striatal circuits. These behavioral problems were assessed in the Frontal Systems Behavior Scale, which questions both the patient and an objective informant (typically, a relative) on symptoms of apathy, disinhibition, and behavioral disorganization. The analyses contrasted the patients' scores with those of the relatives in two time moments: during drug use and during abstinence. The results showed that substance dependents reported significantly lower levels of apathetic and disorganized symptomatology compared with the reports of their relatives. Also, the severity degree of cocaine and alcohol use correlated negatively with the level of discrepancy between addicts and relatives, which suggested an association between the severity of drug use problems and greater levels of lack of awareness on the implications of the addiction.

CONCLUSIONS AND CLINICAL IMPLICATIONS

Deficits in the processes of awareness and denial of addiction and their consequences can have important implications in the treatment and rehabilitation of drug users. During active drug use, the deficits in these processes may be associated with reduced perception of need for treatment, for example, due to lack of insight about the problems or to a tendency to overestimate the own ability to control drug taking without help of others. Furthermore, during rehabilitation, lack of awareness on the own neuropsychological deficits may be associated with reductions in the motivation for treatment, lack of motivation and involvement in conducting the necessary tasks for the achievement of the intervention goals, or a higher feeling of control on the execution of risk behaviors like those entailing the contact with a drug-related context. For this reason, the inclusion of tools aiming to increase the individual's self-awareness during the treatment of addiction could entail significant improvement in the recovering process. These measures could be used to reeducate the interoceptive system leading to an increase of the insight and the perception of body signals anticipating the craving for consuming drugs and potential relapses.

Finally, based on the existing evidence and despite the great efforts made to distinguish the processes' lack of awareness and denial, we still observe a certain overlap between both processes in the neuropsychological studies undertaken in addicted individuals. A more detailed research using specific measure tests of these processes in drug users could help obtain more enlightening information on the involvement of both phenomena in the course of addiction.

SEE ALSO

An Evolutionary Perspective on Addiction, Cognitive Factors in Addictive Processes, Emotions and Addictive Processes, Metacognition in Substance Misuse

Glossary

Confabulation implicit attempts to create an unreal autobiographical story to preserve a personal sense of coherence and continuity in our lives.

Denial implicit suppression of cognitive or emotional information relevant to have adequate knowledge of oneself's current status and of the implications of that status.

Executive functions of higher order abilities involved in the generation, monitoring, and control of complex goal-directed behavior. They are critical to enact and integrate cognitive and emotional input necessary for adequate self-awareness.

Intention–action dissociation lack of correspondence between explicit intention and the translation/implementation of that intention into actual behavior.

Lack of awareness alterations in the access, processing, compiling, or monitoring of cognitive and emotional information critical to have adequate knowledge of oneself's current status and of the implications of that status.

Lie explicit omission of cognitive or emotional information relevant to understand oneself's current status and make it understandable to others. This is done in an attempt to protect oneself from the consequences of truth.

Neuropsychology science investigating the link between brain function and dysfunction and behavior (including cognitions, emotions, and acts).

Self-awareness ability to perceive oneself in relatively objective terms, maintaining a sense of subjectivity. The construct encompasses the cognitive processes necessary for adequate self-knowledge, the emotional processes necessary to understand the implications of the current status for oneself and others, and the integration of cognition and emotion to adjust behavior to one's and others' needs.

Further Reading

Duffy, J.D., 1995. The neurology of alcohol denial: implications for assessment and treatment. Canadian Journal of Psychiatry 40, 257–263.

Ekman, P., 1975. Unmasking the Face: Guide to Recognizing Emotions from Facial Clues. Prentice Hall, Englewood Cliffs, N.J.

Goldstein, R.Z., Parvaz, M.A., Maloney, T., Alia-Klein, N., Woicik, P.A., Telang, F., Wang, G.J., Volkow, N.D., 2008. Compromised sensitivity to monetary reward in current cocaine users: and ERP study. Psychophysiology 45, 705–713.

Goldstein, R.Z., Craig, A.D., Bechara, A., Garavan, H., Childress, A.R., Paulus, M.P., Volkow, N.D., 2009. The neurocircuitry of impaired insight in drug addictions. Trends in Cognitive Sciences 13, 372–380.

Langer, K.G., Padrone, F.J., 1992. Psychotherapeutic treatment of awareness in acute rehabilitation of traumatic brain injury. Neuropsychological Rehabilitation 2, 59–70.

LeDoux, J., Wilson, D.H., Gazzaniga, M., 1977. A divided mind. Annals of Neurology 2, 417–421.

Moeller, S.J., Maloney, T., Parvaz, M.A., Dunning, J.P., Alia-Klein, N., Hajcak, G., Telang, F., Wang, G.J., Volkow, N.D., Goldstein, R.Z., 2009. Enhanced choice for viewing cocaine pictures in cocaine addiction. Biological Psychiatry 66, 169–176.

Moeller, S.J., Maloney, T., Parvaz, M.A., Alia-Klein, N., Woicik, P.A., Telang, F., Wang, G.J., Volkows, N.D., Goldstein, R.Z., 2010. Impaired insight in cocaine addiction: laboratory evidence and effects on cocaine-seeking behaviour. Brain 133, 1483–1484.

Prigatano, G., Weintein, E.A., 1996. Edwin A. Weinstein's contributions to neuropsychological rehabilitation. Neuropsychological Rehabilitation 6, 305–326.

Rinn, W., Desai, N., Rosenblatt, H., Gastfriend, D.R., 2002. Addiction denial and cognitive dysfunction: a preliminary investigation. The Journal of Neuropsychiatry and Clinical Neuroscience 14, 52–57.

Stuss, D.T., Alexander, M.P., 2000. Executive functions and the frontal lobes: a conceptual view. Psychological Research 63, 289–298.

Stuss, D.T., Levine, B., 2002. Adult clinical neuropsychology: lessons from studies of the frontal lobes. Annual Review of Psychology 53, 401–433.

Tomasi, D., Zhang, L., Cottone, L.A., Maloney, T., Telang, F., Caparelli, E.C., Chang, L., Ernst, T., Samaras, D., Squires, N.K., Volkow, N.D., 2007. Is decreased prefrontal cortical sensitivity to monetary reward associated with impaired motivation and self-control in cocaine addiction? American Journal of Psychiatry 164, 43–51.

Verdejo-Garcia, A., Rivas-Pérez, C., Vilar-López, R., Pérez-García, M., 2007. Strategic self-regulation, decision-making and emotion processing in poly-substance abusers in their first year of abstinence. Drug and Alcohol Dependence 86, 139–146.

Verdejo-García, A., Pérez-García, M., 2008. Substance abusers' self-awareness of the neurobehavioral consequences of addiction. Psychiatry Research 158, 172–180.

Relevant Websites

http://www.candaceplattor.com/articles/Denial_and_Addiction.pdf – Definition and types of denial.

http://www.tgorski.com/clin_mod/dmc/dmc.htm – Denial management counseling, T. Gorski.

http://www.proyectohombrenavarra.org/07documentacion/libroilorea.pdf – Brain and addiction, J. Tirapu.

CHAPTER

9

Gateway Hypothesis

Hamdy Fouad Moselhy
UAE University, United Arab Emirates

OUTLINE

INTRODUCTION

Gateway hypothesis describes a pattern (sequences) of drug use, beginning with alcohol and tobacco, followed by cannabis, then more dangerous hard drugs, such as heroin and cocaine. Since the publishing of the original article of gateway hypothesis in *Science* in 1975, a growing body of studies has been pointing to the important role of alcohol, tobacco, and marijuana as gateway drugs. In addition, gateway hypothesis underlies much of the current United States and international drug policy. However, the idea that one behavior leads to another is an old one going back hundreds of years. Development of the gateway hypothesis started with the earlier stepping stone hypothesis. Stepping stone hypothesis was first tested at around 1940 in response to questions about whether marijuana use was the first step toward the use of other drugs. According to the stepping stone hypothesis, the use of one substance or drug increases the probability of consuming another, possibly more harmful, drug later and that the probability increases with frequency of use. It predicts that people who use one substance such as nicotine will probably go on to drink alcohol, to smoke marijuana, and progress on to use hard drugs such as cocaine. Each step is seen as leading to the next step in substance use. Although it might be widely accepted by the public, it has lost credibility and support among research scientists because of the lack of evidence supporting any causality. In contrast, a number of studies tested the gateway hypothesis with findings that supported the gateway sequence. Both longitudinal and cross-sectional studies conducted with general populations in the United States and internationally supported a sequence of drug use progression, beginning with soft licit drugs (both alcohol and cigarettes), followed by illicit drugs (such as marijuana), and ultimately other hard drugs such as heroin and cocaine. One of the most robust findings in adolescent drug use over the past few decades has been that almost all adolescents who have tried cocaine and heroin, first used alcohol, tobacco, and cannabis. The more regularly adolescents use cannabis, and the earlier the age at which they begin, the more likely they are to use other illicit drugs. For example, in one longitudinal study, men who had used both alcohol and cigarettes by age 15 had a 52% greater chance of using marijuana, compared with men who had never used alcohol or cigarettes by age 25. For women, the increased chance of marijuana use among alcohol and cigarette users

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00009-7

was 46%. Similarly, for the next stage, men who had used marijuana by age 15 had 68% greater chance of initiating the use of other illicit substances, compared with those who had never used marijuana. For women, the increased probability was 53%.

Although pharmacological differences exist among the different classes of abused drugs, the one factor that all drugs with a strong dependence potential share is a rewarding or reinforcing property. It is not known why drugs with such different pharmacological properties should share the property of primary reinforcement, and the underlying neuronal mechanisms are not understood. However, several lines of evidence suggest that dopaminergic pathways are implicated in at least some reward circuits and different drugs may activate or "switch on" the circuits at different points.

However, not all research supports the gateway hypothesis. A study of an inner city among New York heavy drug users found that only 33% of drug users followed the gateway hypothesis sequence (i.e. alcohol use to marijuana to hard drug). Another earlier research found that a majority of drug users from the general population were experimental users, and thus relatively few individuals proceeded to regular use of hard drugs (e.g. heroin or cocaine). Most significantly, studies consistently show that individual variations (e.g. genetic makeup, personality traits) and environmental factors (e.g. drug availability, peer influences, life events) may help explain why young people initiate drug use and move on to additional forms of drug use later in life.

There is no question that the sociological, economic, and psychological factors leading to inappropriate use of mood-altering drugs are exceptionally important. However, recent evidence unequivocally establishes the existence of neurobiological determinants of both initial and especially continuing drug uses. All drugs, which have been misused, have one thing in common: they enhance the activity of specific neurobiological circuits. These circuits are commonly described as "brain reward system." The most prominent neurotransmitters and neuropeptides in this process are dopamine, serotonin, glutamate, gamma-aminobutyric acid, and the endorphines (endogenous opioids). Dopamine appears to be playing the primary role in producing most of euphoria seen with drug of abuse and is known to contribute to the positive reinforcing effects of heroin and other opiates.

The Gateway Pattern of Drug Use

The gateway pattern describes the typical sequence of progression to hard drug use. The sequence most often reported is that alcohol and tobacco use come first, followed by marijuana, and then other illicit substances. Obviously, several factors affect such a sequence, including drug availability and background prevalence. Some illicit drug use is significantly more common among more recent birth cohorts and the pattern of availability has changed. Possibly, it is not astonishing that a "reverse gateway" has been described for cannabis in Australia (where cannabis use has been linked to increased risk of subsequent initiation to tobacco use). Some variations in the typical sequence have been found for individuals of different sexes, racial and ethnic groups, and cultures. There have been investigations of the extent and significance of violations of typical gateway patterns. Studies in the United States of problematic drug users and homeless youths have found that significant proportions had not progressed through the typical pattern of progression, with many beginning cannabis use before they had first used alcohol, and some starting other illicit drug use before using alcohol or cannabis. In those studies, individuals with "atypical" patterns of progression were found to come from more disadvantaged backgrounds, and also, were found to be heavier polydrug users than users who followed the normative progression. However, those studies presented limited unrepresentative samples of heavy drug users and did not adjust for premorbid mental health or demographic factors that might have been related to progression. Another study was done that considered all of these confounding possibilities using data from a representative sample of the US adult population, from the National Comorbidity Survey Replication. The study showed that deviations from the gateway order of onset were found to occur only for a minority of persons (5.2%). The most common pattern was other illicit drug use before cannabis (3.4%), and the least common was other illicit drugs use before both alcohol and tobacco use (0.8%). There were some strong cohort differences in the likelihood of these patterns: they were less common among the oldest age group than the younger ones. These findings are consistent with historical trends in drug use; cannabis use is much more common in more recent birth cohorts, so it is not surprising that cannabis is also more likely to occur earlier in the sequence of drug use for some younger people. One significant predictor of deviations from normal pattern was the early development of internalizing mental disorders such as depression, posttraumatic stress disorder, social phobia, or generalized anxiety disorder. The authors suggested that premorbid mental illnesses are related to precocious initiation of illicit drug use. Adolescents who already developed mental illness are at risk for deviations from the normative sequence of drug initiation and for the development of dependence. However, this study was cross-sectional retrospective survey with some limitations including, some of the observed cohort differences might be traced

to higher mortality among individuals in the older cohorts who began drug use at an early age. Additionally, retrospective reporting of the age of first drug use may be subject to error, given that respondents are being asked about events that may have occurred decades ago. Finally, patterns of progression from one drug to another are not, however, consistent or predictable for all individuals. They can vary widely, depending on numerous individual and contextual factors and on characteristics of the use itself, such as age of onset, type of drug, frequency, and quantity of substance used.

THE MECHANISMS OF PROGRESSION IN DRUG USE

Soft drug abuse can be an important contributor to later hard drug abuse, by affecting an individual physiologically, psychologically, and socioculturally.

Physiological Mechanisms

Some authors have speculated that the pathway from use of alcohol and tobacco to illicit drugs may be facilitated by effects of early-stage drug use on central reward circuitry. Such suggestions are based on observations in animal models that exposure to some drugs alters future behavioral responses to the same or other drugs. Neuroanatomical, neurochemical, and molecular studies in animals show that nicotine exposure during adolescence produces a sensitization to drugs of abuse in young adulthood. So far, the mechanisms reported due to nicotine administration include alteration in the dendritic morphology of medium spiny neurons within the nucleus accumbens; reduced choline acetyl tranferase activity and increased hemicholinium-3 binding within the midbrain; increased dopamine transporter densities and decreased serotonin transporter densities; reduced DNA content within the midbrain, cerebral cortex and hippocampus; and an increased expression of specific nicotinic acetylcholine receptor subunits within the ventral tegmental area. The persistent neurobiological effects of periadolescent nicotine priming within structures associated with the reward pathway may translate into a sensitization to various drugs of abuse by the production of an increased response within the reward pathway itself. The potential for an increased response within the reward pathway may be mediated through localized alterations in neurotransmitter systems, namely dopamine, within the ventral tegmental areas, nucleus accumbens, or prefrontal cortex.

Addictive substances affect processes in the brain operating through reward systems and can produce drug tolerance and dependence. Tolerance manifests itself when, to produce a given response (e.g. a high), an individual must ingest more of a substance. It is the state in which repletion of the same dose of a drug has progressively less effect or in which the dose needs to be increased to obtain the same degree of pharmacological effect as was caused by the original dose. Although tolerance itself does not directly increase the likelihood of continued or compulsive use of drugs, such as narcotics or CNS depressants, it can affect the pattern of abuse by increasing the dose to experience the reinforcing effects. Use of increased amounts in turn enhances the risk of toxic effects and problems (including expenses) connected with securing the drugs.

The mechanisms involving the development of tolerance are only partially understood. In animals, tolerance often occurs as a result of induction of hepatic microsomal enzyme synthesis concerned with drug biotransformation (drug-disposition or pharmacokinetic tolerance) or of neuronal adaptation to drug action "pharmacodynamic, tissue, or cellular" tolerance. Further mechanisms of tolerance have been investigated at the level of "receptors," or at the molecular levels concerning synthesis and metabolism of protein or various neurotransmitters.

Physical dependence is viewed as one of several factors that contribute to the development of, and to the tendency to relapse after, withdrawal. It has been suggested that withdrawal symptoms associated with several classes of drugs of abuse are characterized by rebound effects in the same physiological systems that were initially modified by the drug (rebound hyperexcitability). Many theories have been proposed to explain the phenomena of physical dependence. However, at present it is difficult to explain by any single model all the complex phenomena associated with different types of drugs causing tolerance and dependence. Involvement of multiple mechanisms is very likely in these phenomena.

The development of physiological tolerance and dependence can contribute to the progression from one drug to another drug. Repeated use of different drugs may lead to the development of two phenomena: cross-tolerance and cross-dependence. Cross-tolerance is defined as "the state in which tolerance to one drug has the effect of causing tolerance to another drug of the same or a different chemical type." On the other hand, cross-dependence is the ability of one drug to suppress the withdrawal manifestations of another and to maintain the state of physical dependence produced by the latter. This phenomenon can be complete or partial. Cross-dependence is generally seen between the potent opioids. Animal studies also show a high degree of cross-dependence among general CNS depressants. In humans, partial cross-dependence is seen between alcohol and barbiturates.

Another important phenomenon in the physiological mechanism of polydrug use is "drug as a reinforcer." Certain drugs when administered create a psychological condition, which the user would like to experience again by repeating the drug use. Such a condition may be a feeling of pleasure, relief of pain and discomfort, or even an altered perception. Such a reinforcing effect that is experienced after self-administration of a drug by its abuser has also been demonstrated in experimental animals such as rats, monkeys, and baboons in the laboratory environment.

Finally, interactions of different classes of drugs have been proposed to support the physiological mechanism of gateway hypothesis. Historically, cannabinoids were used in combination with opioids for the treatment of different types of pain in humans due to their synergistic interactions in the modulation of noxious stimuli. Cannabinoids produce a variety of pharmacological effects very similar to those elicited by opioids. Direct and indirect interactions with opioid system have been proposed to explain some cannabinoid effects such as analgesia and attenuation of opioid-withdrawal syndrome. Evidence has been found in support of the notion that rewarding properties of cannabinoids and opioids might be functionally linked. In particular, a growing body of studies points to an important role of endogenous cannabinoid system in the modulation of opioid rewarding and addictive effects. It is generally acceptable that the use of cannabinoids is related to their positive modulatory effects on brain-rewarding processes along with their ability to positively influence emotional states and remove stress responses to environmental stimuli. Cannabinoids have been tested on a variety of behavioral models of addiction, most of which revealed functional interactions between the endocannabinoids and opioid systems in the modulation of reciprocal rewarding and addictive effects.

Cannabinoids and opioids share many pharmacological properties, including antinoception, hypothermia, sedation, and inhibition of intestinal motility. Chronic administrations of both agents produce tolerance to their analgesic and hypothermic effects and lead to the development of physical dependence, although with different intensities. Besides analgesia, endogenous cannabinoids interact with the opioid system in a variety of biological functions, including emesis, intestinal motility, and immune activity as well as modulation of anxiety, stress, emotion, and exploratory behavior. When the caudate-putamen (CP) of rats are treated with repeated administration of central cannabinoids (CB1) receptor ligand, delta 9-tetrahydrocannabnoid, there is an increase in proenkephalin gene expression and u-opioid receptor activation of G-proteins, a time-related decrease in central cannabinoids (CB1) receptor gene expression, and a reduction in CB1 receptor activation of G-proteins. These findings suggest a possible interaction between the cannabinoids and opioid systems in a brain area (i.e. caudate-putamen).

These mechanisms are consistent with "priming effect" which was suggested by researchers. Priming effect is the effect in which exposure to a stimulus repetition, perceptual or conceptual, influences response to a later stimulus. Animal experiments in addiction showed that after a period of drug self-administration followed by a period of abstinence, resumption of self-administration occurs much more readily than in animals never exposed to the drug. This priming effect of the first small dose after prolonged abstinence is thought to play an important part in provoking full-blown relapse. Many research projects suggest by the speculation that the conditioned cue causes "priming" by release of an endogenous activator of the reward pathway. Drug-related cues evoke conditioned drug-like effects which can "prime" a return to drug taking. The priming effect can readily explain why drug-related cues tend to evoke craving in drugs, without appealing to conditioned withdrawal symptoms. However, very few animal experiment have been conducted in relation to gateway hypothesis, and most have focused on priming one class of drug on the subsequent use of the same drug.

Psychological Mechanisms

The use of soft drugs may contribute to further abuse of harder drugs through psychological processes as well. For example, an initial successful experience of use may reduce an individual's fear about the first drug, thus opening the way to continued use and progressing to use other drugs. Additionally, continued use of substances can also impair the learning processes including the ability to remember lessons learned in earlier stages. Furthermore, developmental processes may be slowed, which may obstruct decision making about the use of drugs. The personality profile of the abuser may play important roles in initiation and maintenance of drug abuse. Individuals with certain personality deficits (inadequate, immature, rather unstable person with a low tolerance for frustration) may be more vulnerable to drug abuse. However, most people with such personality traits do not become chronic drug users unless other predisposing factors are present. But such individuals are vulnerable to these predisposing factors. They become motivated to have drug-induced experiences, particularly those providing pleasure, relief of pain, and altered perceptions.

One factor that can influence whether youths will use tobacco, alcohol, or illicit drugs is the extent to which youths believe these substances might cause them harm. National Survey of Drug Use and Health (NSDUH) in

the United States, 2008 respondents were asked how much they thought people risk harming themselves physically and in other ways when they use various substances in certain amounts or frequencies. Response choices for these items were "great risk," "moderate risk," "slight risk," or "no risk." The percentages of youths reporting binge alcohol use and use of cigarettes and marijuana in the past month were lower among those who perceived great risk in using these substances compared with those who did not believe as great risk. The results show that 5.0% of youths aged 12–17 who perceived great risk from "having five or more drinks of an alcoholic beverage once or twice a week" reported binge drinking in the past month (consumption of five or more drinks of an alcoholic beverage on a single occasion on at least 1 day in the past 30 days); in contrast, past month binge drinking was reported by 11.5% of youths who saw moderate, slight, or no risk from having five or more drinks of an alcoholic beverage once or twice a week. Past month marijuana use was reported by 1.5% of youths who saw great risk in smoking marijuana once a month compared with 9.4% of youths who saw moderate, slight, or no risk.

Sociocultural Mechanisms

Several hypotheses have focused on either personality characteristics or environmental circumstances in the use of drugs and the progression of drug abuse. Sensation-seeking traits have been hypothesized to predict the initial incursion into testing with drugs. Because of the relative ease of obtaining nicotine and alcohol, these drugs would be most likely to be used first. Initial use of other illicit drugs may follow as a result of continuing search for new sensations and exposure to environmental conditions that are more willing to obtaining these drugs. Although genetic variability accounts for some amount of drug use behaviors and likelihood of addiction, there are other individual factors related to drug use such as deviance proneness and personality variables (e.g. impulsivity, negative emotionality, rebelliousness, low self-esteem, poor scholastic achievement, truancy, decreased academic aspirations, and lack of motivation).

Furthermore, an individual who initiates use may begin to participate in a subgroup that encourages use, such as adolescent peer groups. Peer group is a very strong shaper of behavior, including drug-taking behavior. Numerous research findings reveal the influence of peer groups as the most dominant factor for an adolescent's initiation and recent use of drugs. Other group level factors relevant for drug use among adolescents include parental attitudes and behaviors, and lifestyle patterns (e.g. frequently going out at night).

Community level factors relevant for drug use behaviors include drug availability in neighborhoods and the popularity of specific drugs. Such social and cultural environment encourages, reinforces, maintains, and increases substance use. Alcohol and nicotine are legal drugs in many countries. Even though these drugs are "socially acceptable" for adults, they are, to one degree or another, controlled substances for youth throughout the world. However, smoking teaches drug acquisition skills to the youngsters. In addition, children who smoke get firsthand experience in using a substance to adjust their emotional states.

The social environment of an abuser, where drug abuse is accepted and prevalent among friends and associates may be one of the most important factors in inducing and maintaining the drug use. The environment with its settings and rituals may serve as a secondary reinforcement. Sometimes such social reinforcement maintains drug use behavior in the initial experimental stage until the newcomer begins to appreciate the primary drug effect or becomes tolerant to some initial aversive effects of a particular drug. For example, some young people may not like the initial effects of smoking tobacco, or may not experience anything pleasurable in smoking marijuana, or may experience nausea and vomiting with an initial dose of heroin; the social reinforcers may maintain the drug use behavior until these initial difficulties are conquered.

Experimental studies with animals and humans have corroborated the secondary reinforcing effects of environment. Acute drug effects, withdrawal manifestations, and their relief by drugs have been shown to be conditioned to environmental stimuli. These phenomena also play important roles in the relapse of drug addiction. Thus a former narcotic dependent may feel craving for the drug when he or she returns to a familiar environment; an alcoholic may have a similar experience when exposed to the sight and smell of alcohol.

Several theories may explain the deviance behavior. First, differential association theory postulates that deviance arises when the values of membership in one group conflict with the values of more powerful groups who are able to embody their definitions into law. People engage in acts defined as deviant because of the values of their reference groups, not because of any abnormal processes. Using the social learning model, researchers demonstrate that the differential association variables explain the most variance in both alcohol and drug use among adolescents. In agreement with prior research, the strongest predictor of substance use is having friends who use drugs and alcohol. Second, control theory posits that deviance arises when young people lack sufficient ties to conventional social groups such as

families, schools, and churches. From this point of view, youths deviate not through frustrated desires or actions in accordance with their own reference groups but because their ties to conventional groups are broken or underdeveloped. Finally, strain theory emphasizes entirely different processes than either differential association or control theory. Its central thesis is that people become prone to deviation when society is unable to satisfy their fundamental needs. A broad version of strain theory views various forms of deviance as mechanisms that allow people to cope with the stresses of everyday life.

Meanwhile, the problem behavior theory has shown that drug abuse is one of a cluster of problem behaviors that also include truancy, delinquency, unhealthy eating habits, excessive TV watching, reckless driving, and premature or reckless sexual behavior. These behaviors do not cause one another; rather, they are common manifestations of traumatized and aimless lives. Moreover, some young people are more vulnerable to destructive habits than others. They are more vulnerable because of psychological maladjustment, family disruption, and economic and social deprivation. So problem behavior theory suggests that what may explain progression from one drug to the next is not only the nature of the first drug used, but also factors in the various systems of which a person is part. That is, individuals, group, community factors interact to increase the risks of rebellious behavior including drug abuse. At the same time, the integrated system theory suggests that numerous factors at various levels of analysis – from cell to society – affect the likelihood of deviant behavior, including drug abuse.

AN ASSESSMENT OF THE VALIDITY OF GATEWAY HYPOTHESIS

Irrespective of the mechanism of association, several studies support the gateway pattern in drug addiction. In the Australian adolescents study, a random sample of 1943 adolescents was recruited from secondary school age 14–15. The participants were interviewed on eight occasions until the age of 24–25 years. The strongest predictor of use at age 24 was the use of other drugs, particularly cannabis at age 20. Furthermore, in the twin study of Virginia Commonwealth University, they found that cannabis was strongly predictive of use of other drugs in the future.

Based on national data that look at both children and adults and all gateway drugs, using 1991 National Household Survey, the Centre on Addiction and Substance Abuse at Columbia suggests that the early the use of gateway drugs, tobacco, alcohol, or cannabis, the more likely they are to move on to other drugs. Youth who drank alcohol were 50 times more likely to use cocaine, and those who smoked tobacco cigarettes were 19 times more likely to use cocaine. Nearly 90% of cocaine users had smoked tobacco, drank alcohol, or used marijuana first. Adults who drank as children are six times more likely to be regular cocaine users. This study, based on 30 000 households, established a clear progression that began with use of gateway drugs of alcohol, tobacco, or cannabis and led to use of other drugs. Using data from a cross-sectional survey in California, 11 239 subjects (46.3% male) from 31 high schools entered the study. After controlling for seven variables, the risk ratio of the last 30-day alcohol use among prior smoking initiators versus noninitiators was 5.82 for non-Hispanic whites, 4.25 for blacks, 8.37 for Asian Indians, 3.99 for Chinese, 3.45 for Filipinos, 3.48 for Japanese, 5.41 for Koreans, 7.57 for Vietnamese, 4.02 for Mexicans, 2.44 for South/Central Americans, and 5.95 for adolescents with multiethnic background. Comparison of the 11 ethnic groups indicated that adolescents from different ethnic groups but with similar cultural background had similar risk level; such a pattern existed after controlling for acculturation, parents' monitoring, and school performance. The risk ratio did not differ by gender and grade. There is an association between prior cigarette smoking initiation and current alcohol use among adolescents from different ethnic backgrounds, including those of multiethnicity, which supports the generalizability of gateway drug effect of cigarette smoking on alcohol use. NSDUH, 2005, found that compared with lifetime nondrinkers, adults who have consumed alcohol were statistically much more likely to currently use illicit drugs or abuse prescription drugs in the past year. Effects were strongest for cocaine (26 times more likely) and cannabis (14 times more likely). Furthermore, lifetime drinkers were six times more likely to abuse or be dependent on illicit drugs than lifetime nondrinkers. Use of illicit drugs and alcohol was more common among current cigarette smokers than among nonsmokers in 2005, as in 2002 through 2004. Among persons aged 12 or older, 20.2%, of past month cigarette smokers, reported current use of an illicit drug compared with 4.1% of persons who were not current cigarette smokers. Past month alcohol use was reported by 67.6% of current cigarette smokers compared with 46.6% of those who did not use cigarettes in the past month. The association also was found with binge drinking (43.8% of current cigarette users vs. 15.7% of current nonusers) and heavy drinking (16.1 vs. 3.5%, respectively). More recent research has supported the gateway sequence. For example, one study of thousands of adolescents found that marijuana users were twice as likely as nonusers to use illicit drugs as young adults. Although shared environmental factors mediated

much of the relationship between adolescent marijuana use and drug use as a young adult, the association remained even when controlling for familial and other factors. Another study, using data from the National Youth Survey, could not discount the possible casual influence of marijuana on other illicit drug use even after controlling for factors predicted by strain, social bonding, and differential association theories. Furthermore, another study published in 2008, examined the applicability of the gateway hypothesis to drug use patterns of secondary school students from a nonmetropolitan area in Tennessee. The data were collected from students in the 8th, 10th, and 12th grades at three secondary schools, using self-administered questionnaires under the supervision of teachers. Although there is some support for the gateway hypothesis in the data, there is also evidence that what differentiates those who move from initial marijuana use to use of harder drugs have risk factors unique to individuals and their environments, consistent with the predictions of problem behavior theory and integrated system theory. Finally, in a 2008 survey, the rate of current illicit drug use was more than nine times higher among youths aged 12–17 who smoked cigarettes in the past month (49.0%) than it was among youths who did not smoke cigarettes in the past month (5.3%). Past month illicit drug use also was associated with the level of past month alcohol use. Among youths aged 12–17 in 2008 who were heavy drinkers (i.e. consumed five or more drinks on the same occasion on each of the 5 or more days in the past 30 days), 68.5% also were current illicit drug users, which was higher than the rate among nondrinkers (4.3%). The rate of current illicit drug use among youths reporting heavy drinking in the past month increased from 60.1% in 2007 to 68.5% in 2008, and a similar increase in illicit drug use (from 37.9 to 42.6%) was seen among youths who engaged in binge drinking (i.e. consumption of five or more drinks on the same occasion on at least 1 day in the past month).

Preclinical studies added more evidence for the gateway hypothesis, in that the Karolinska Institute in Sweden used rats to investigate the effect of cannabis use and the subsequent use of other drugs in the future. The findings supported the gateway hypothesis showing that adolescence cannabis exposure has a lasting effect on hedonic processing resulting in enhanced opioid intake, possibly through the alteration of limbic opioid neurons. On examining the brain cells of the rats, they found that tetrahydrocannabinoids (THC) alters the opioid system that is associated with positive emotions, which decreases the effects of opioids on rat's brain and caused them to use more heroin.

Animal studies results establish a relationship between nicotine use by adolescents and a subsequent involvement with drugs of abuse in adulthood. The rodent model of periadolescence correlates strongly with the behavioral parameters that characterize human periadolescence: periadolescent rodents display increased impulsivity, amplified novelty seeking, diminished novelty-induced anxiety, and increased risk-taking. Exposure to nicotine during the periadolescent period affects integrated learning in the postadolescent rat by enhancing performance on the passive avoidance test in the posttreatment period. Importantly, periadolescent nicotine priming produced chemical alterations in anatomical structures comprising the reward pathway. Furthermore, periadolescent exposure to nicotine increase dopaminergic turnover in the young adult rat within the midbrain. Moreover, periadolescent nicotine priming was shown to increase young adult animal's sensitivity to the reinforcement of cocaine in the conditioned place preference paradigm of reward and reinforcement. Finally, nicotine conditioning produced a dose-dependent place preference in the periadolescent rat but did not produce a place preference in the postadolescent rat. Another study shows that rats treated with only nicotine during periadolescence showed a sensitization to the subthreshold dose of diazepam utilized during conditioning. These studies support the role of nicotine in the production of a gateway drug effect as periadolescent nicotine priming has shown to produce sensitization to nicotine itself, cocaine, and diazepam. Finally, the periadolescent period appears to be a critical time in the development of the neural system and encompasses a time of neurobehavioral vulnerability to gateway drugs and the production of persistent, behavioral adaptation in the young adult.

CHALLENGING ACCOUNTS OF THE GATEWAY HYPOTHESIS

While some research discussed earlier support the gateway hypothesis, other research challenges these findings. In the Harvard alcohol survey, 7 in 10 students had drunk in the last month. But according to a 1994 government survey, only 12% of 18–25 years old and 6% of 12–17 years old had taken marijuana during past month, only 1% of 18–25 years old had taken cocaine, and far fewer had taken crack or heroin. The study of "Marijuana and Medicine: Assessing the Science Base," in 1999, by the Division of Neuroscience and Behavioral Health at the Institute of Medicine found no evidence of a link between cannabis use and subsequent abuse of other illicit drugs on the basis of its particular physiological effect. In 2002, a study by RAND Drug Policy Research Centre casts doubt on claims that marijuana acts as a "gateway" to the use of cocaine and heroin. RAND researchers tested the marijuana gateway theory

by creating a mathematical model simulating adolescent drug use. Rates of marijuana and hard drug use in the model matched those observed in survey data collected from representative samples of youths from across the United States. Without assuming any gateway effect, the model produced patterns of drug use and abuse remarkably similar to what is experienced across the nation, showing that the marijuana gateway effect is not needed to explain the observed behavior. The association can result from known differences in the ages at which youths have opportunities to use marijuana and hard drugs, and known in individuals' willingness to try any drugs. In 2004, a study comparing cannabis users in San Francisco to those in Amsterdam was done to test the effects of the differing drug policies in the two cities on drug use patterns. The Netherlands has a drug policy of decriminalization in which cannabis can be bought by adults over 18 in quasi-legal coffee shops and used publicly, while in the United States cannabis is criminalized and must be bought in black market and used secretly. The results found that the San Francisco cannabis users were significantly more likely to use cocaine, crack, amphetamine, ecstasy, and opiates despite similar cannabis use patterns and a more permissive drug policy in the Netherlands. One explanation is that the black market itself acts as a gateway to harder drugs, as opposed to the effects of cannabis per se. In 2006, a 12-year gateway drug hypothesis study on 214 boys aged 10–12 by the American Psychiatric Association concluded that adolescents who used cannabis before using other drugs, including alcohol and tobacco, were more likely to develop a substance abuse disorder than subjects in the study who did not use cannabis prior to using other drugs. This means that changing the order of gateway drugs did not change the outcomes. Finally, new research from the University of New Hampshire, published in September, 2010, shows that the gateway effect of marijuana is overstated. The researchers used survey data from 1286 young adults who attended Miami-Dade public schools in 1990s. Within the final sample, 26% of the respondents are African American, 44% are Hispanic, and 30% are non-Hispanic white. The researchers found that young adults who did not graduate from high school or attend college were more likely to have used marijuana as teenagers and other illicit substances in young adulthood. In addition, those who used marijuana as teenagers and were unemployed after high school were more likely to use other illicit drugs. However, the association between teenage marijuana use and other illicit drug abuse by young adults fades once stresses, such as unemployment, diminish. In addition, once young adults reach age 21, the gateway effect subsides entirely.

Research from animal studies showed that cannabis derivatives have proved to exert highly specific drug discrimination effects, which are not substituted by other classes of drugs (i.e. opioids) nor are they reversed by antagonists of various neurotransmission systems, supporting the idea that the drug discrimination effects only involve the cannabinoid system. Accordingly, morphine does not substitute for THC. In addition, the hypothesis that prior cannabis exposure increases the likelihood of becoming addicted to other drugs was evaluated by giving rats a THC, then allowing them to self-administer other drugs. THC exposure did not alter the acquisition of cocaine self-administration or the amount of cocaine taken under a fixed-ratio schedule.

SUMMARY

The sequence of drug use is far from clear and any discussion of this topic generates much controversy. According to the gateway hypothesis, there is a progressive and hierarchical sequence of stages of drug use that begins with tobacco or alcohol and proceeds to marijuana, and from marijuana to other illicit drugs, such as cocaine and heroin. The basic idea of the hypothesis is that taking part in different classes of drugs is not opportunistic but follows definite pathways; an individual who participates in one drug is at risk of progressing to another. Evidence from research has shown that there does appear to be an agreed-on model of sequential stages of involvement with substances, which typically begins with alcohol. In addition, there is a clear association between frequency of use of a lower-stage drugs and likelihood of initiating high-stage drugs. However, this conclusion can be challenged on the grounds of having several confounding factors in most of these studies, which can include genetic, environmental, and individual variables. A recent study found that the strongest predictor of whether someone will use other illicit drugs is their race/ethnicity, not whether they ever used marijuana. Also, teenagers who are stressed and unemployed after graduation from high school were more likely to use other illicit drugs. Although most hard drug users were soft drug users, there is no conclusive evidence that the gateway drugs are causally linked to subsequent abuse of other illicit drugs as association does not mean causality. In addition, the link between marijuana and other illegal drugs stems from the fact that they are illegal and only available in the black market. Young people who go into the drug market are more at risk to be exposed to drug users and drug dealers and so urged to try more drugs. In conclusion, the pattern of drug addiction and its pathways are likely to be both multifaceted and complex. There is likely to be an intricate

relationship between pre-existing neuropsychological vulnerabilities, the age of initiation of substance use, patterns of substance use (type, dosage, and duration), availability of the drugs, and associated adverse events.

SEE ALSO

An Evolutionary Perspective on Addiction, Tolerance and Withdrawal, Implicit and Associative Processes in Addiction

Glossary

Cannabis cannabis products made entirely of plants in the genus *Cannabis sativa*. It is a weedy annual that grows abundantly in the tropical and temperate zones of the world. It is the sole species containing the psychoactive substances (cannabinoids) and is of two varieties: *C. sativa indica* and *C. sativa americana*.

Gateway drug a drug (as alcohol or marijuana) whose use is thought to lead to the use of and dependence on a harder drug (as cocaine or heroin).

Gateway hypothesis gateway hypothesis describes a pattern (sequences) of drug use, beginning with alcohol and tobacco, followed by cannabis, then more dangerous hard drugs, for example, heroin and cocaine.

Hard drugs a psychoactive drug that is considered relatively strong, likely to cause addiction and perceived as specially damaging such as heroin and crack cocaine. The terms "hard" and "soft" when applied to drugs have no legal or pharmacological validity.

Priming effect effect in which exposure to a stimulus repetition, perceptual or conceptual, influences response to a later stimulus. Animal experiments in addiction showed that after a period of drug self-administration followed by a period of abstinence, resumption of self-administration occurs much more readily than in animals never exposed to the drug. This "priming effect" of the first small dose after prolonged abstinence is thought to play an important part in provoking full-blown relapse. Many research projects are suggested by the speculation that the conditioned cue causes "priming" by release of an endogenous activator of the reward pathway. drug-related cues evoke conditioned drug-like effects that can "prime" a return to drug taking. The priming effect can readily explain why drug-related cues tend to evoke craving in drugs, without appealing to conditioned withdrawal symptoms.

Soft drug a drug of abuse that is considered relatively mild and not likely to cause addiction such as cannabis and LSD. The term "soft drug" is considered controversial by its critics because it implies that the drug causes no or insignificant harm.

Stepping stone hypothesis the use of one substance or drug increases the probability of consuming another drug later and that the probability increases with frequency of use.

Tobacco tobacco products are products made entirely or partly of leaves of plants of the genus *Nicotiana tobacco*. They are used as raw material, which are intended to be smoked, sucked, chewed, or snuffed. All contain the highly addictive psychoactive ingredient, nicotine.

Further Reading

Chen, X., Unger, J., Palmer, P., Weiner, M., Johnson, C., Wong, M., Austin, G., 2002. Prior cigarette smoking initiation predicting current alcohol use: evidence for a gateway drug effect among California adolescents from eleven ethnic groups. Addictive Behaviours 27, 799–817.

Choo, T., Roh, S., Robinson, M., 2008. Assessing the gateway hypothesis among middle and high school students in Tennessee. The Journal of Drug Issues 0022-0426/08/02, 467–492.

Degenhardt, L., Chiu, W.T., Conway, K., Dierker, L., Glantz, M., Kalaydijian, K., Merikangas, K., Sampson, N., Swendsen, Kessler, R.C., 2009. Does the gateway matter? Associations between the order of drug initiation and development of drug dependence in the national comorbidity study replication. Psychological Medicine 39, 157–167.

Fattore, L., Deiana, S., Spano, S.M., Cossu, G., Fadda, P., Scherma, M., Fratta, W., 2005. Endocannabinoid system and opioid addiction: behavioural aspects. Pharmacology Biochemistry and Behavior 81, 343–359.

Fergusson, D.M., Boden, J.M., Horwood, L.J., 2006. Cannabis use and other illicit drug use: testing the cannabis gateway hypothesis. Addiction 101, 556–569.

Kandel, D.B., 1975. Stages in adolescent involvement in drug use. Science 190, 912–914.

Kandel, D., 2002. Stages and Pathways of Drug Involvement: Examining Gateway Hypothesis. Cambridge University Press, Cambridge, England.

Kandel, D., 2003. Does marijuana use cause the use of other drugs? JAMA 289, 482–483.

Moral, A.R., McCaffrey, D.F., Paddock, S.M., 2002. Reassessing the marijuana gateway effect. Addiction 97, 1493–1504.

Peele, S., 1998. Gateway to nowhere: how alcohol came to be scapegoated for drug abuse. Addiction research 5, 419–426.

Stewart, J., 1984. Reinstatement of heroin and cocaine self-administration behavior in the rate by intracerebral application of morphine in the ventral tegmental area. Pharmacology Biochemistry and Behavior 20, 917–923.

Relevant Websites

http://www.peele.net/lib/gateway.html – Gateway to Nowhere.

http://www.druglibrary.net/schaffer/Library/studies/ota/ch5.htm – The Schaffer Library of Drug Policy.

http://www.suite101.com/content/gateway-drugs-a17250#ixzz1BqlZ4Ow0 – Looking at the Definition, Theories, Controversy: Drug Gateway.

http://www.sciencedirect.com/ – Evaluating the Drug Use Gateway Theory.

http://www.physorg.com/news202619339.html – Risk of Marijuana's Gateway Effect Overblown.

http://www.rand.org/news/press.02/gateway.html – Rand Study.

http://clearinghouse.missouriwestern.edu/manuscripts/481.php – Is Alcohol a Gateway Drug.

http://www.drugwatch.org/alcohol%20&;%20Tobacco_Gateway%20Drus.htm – Alcohol and Tobacco: Two Dangerous Gateway Drugs.

http://www.informaworld.com/smpp/content – The Effectiveness of Gateway Communications in Anti-Marijuana Campaigns.

http://www.mjlegal.org/gateway.html – Marijuana and the Gateway Theory.

CHAPTER

10

Adolescent Substance Use: Symptoms and Course

Tammy Chung

University of Pittsburgh Medical Center, Pittsburgh, PA, USA

OUTLINE

ADDICTION: A DEVELOPMENTAL DISORDER

Addiction is a developmental disorder that often begins in adolescence. The adolescent developmental period, which generally covers ages 12–18, involves rapid changes in physical, cognitive, and affective domains that contribute to a drive for greater independence, as well as to a normative increase in risk-taking behavior, such as substance use. These adolescent-specific changes converge to make adolescence a period of peak risk for substance use onset and provide the rationale for using a developmental approach to understanding addiction. A developmental approach proposes that multiple systems (e.g. individual, family, school, neighborhood, country),

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00010-3

operating at multiple levels (e.g. individuals nested within families that are contained within neighborhoods), interact to influence an adolescent's trajectory of substance use. In this approach, risk factors (e.g. peer drug use) interact with protective factors (e.g. high parental supervision) dynamically over time to influence trajectories of substance use. Most risk and protective factors for substance use are common across substances; substance-specific risks generally involve processes related to drug-specific metabolism. Differences between adolescent and adult substance users with regard to, for example, legality of alcohol and tobacco use, contexts and reasons for use, duration and severity of use, and type of substance-related problems most often experienced, also point to the importance of using a developmental approach to understanding adolescent substance involvement.

Substance use ranges along a continuum from no lifetime use, experimentation, regular use, and heavy use through addiction. Addiction refers to a pattern of compulsive substance use, often despite significant harm or impairment in functioning caused by substance use. Progression of substance involvement is probabilistic, not inevitable. Only a minority of individuals who initiate use of a substance progress to addiction, and among substance users, periods of escalation in use may alternate with periods of reduction in use and desistance. Trajectories of substance use in adolescence typically represent developing (rather than fully formed) entities or phenotypes, only some of which eventuate in addiction. The following sections review adolescent-specific maturational processes that contribute to risk for substance use, externalizing and internalizing behavior pathways of risk, the sequence of substance use initiation, the prevalence of adolescent substance use, contexts and motives for use, prototypical trajectories of adolescent substance use, the prevalence of substance use disorders (SUDs) in youth, and the course of adolescent substance-related problems. A final section reviews implications for intervening with youth.

ADOLESCENT DEVELOPMENTAL MATURATION AND RISK FOR SUBSTANCE INVOLVEMENT

Developmental changes specific to adolescence that have been associated with risk for substance use include early pubertal maturation, continuing brain development, and increased sensitivity to drug effects. Early pubertal maturation increases risk for adolescent substance use because precocious physical development can facilitate association with older peers, who might provide access to drugs and socialize youth to substance use. In addition, among early maturers, cognitive and social skills often lag behind physical development, potentially resulting in situations in which an early maturer might lack the interpersonal skills and cognitive control needed to refrain from engaging in risky behavior, such as substance use. Continuing brain development during adolescence, which occurs independent of pubertal maturation, involves earlier maturation of neural systems associated with emotion and reward processing, relative to cognitive control systems. The later maturation of cognitive control, relative to emotion processing, is thought to underlie the normative adolescent increase in risk-taking behavior and contributes to the developmentally specific increase in risk for onset of substance use during adolescence.

Another biologically based developmental risk factor involves adolescents' greater sensitivity to the positive or rewarding effects of alcohol and other drugs, compared with adults. For example, an animal model research suggests that adolescents, compared with adults, are more sensitive to alcohol's stimulant and social facilitation effects. In addition, adolescent animals are less sensitive to alcohol's sedative effects, which serve as cues to limit intake. Adolescents' greater sensitivity to "positive" alcohol effects, coupled with decreased sensitivity to "negative" effects, contributes to adolescents' propensity to engage in episodes of high-volume consumption, a particularly risky pattern of drinking. Likewise, adolescent rats self-administer nicotine more often and in higher doses than rats initially exposed as adults. Individual differences in initial sensitivity to a drug may further augment adolescent-specific sensitivity, resulting in a "double-dose" of vulnerability for certain youth. Another possible effect of substance use during adolescence involves "sensitization." The initiation and use of certain substances, particularly during adolescence, can result in sensitization to the substance, that is, potentially long-lasting changes (e.g. substance-induced changes in gene expression) that increase risk for heavy use.

These biologically based developmental changes, which are specific to the adolescent developmental period, occur in the context of a social environment that provides increasing opportunities to engage in substance use (e.g. decreased parental monitoring, greater access to substances). Importantly, individuals actively seek out and participate in the creation and maintenance of the social environments that they inhabit. Thus, interventions that aim to reduce risk for adolescent substance use need to address the adolescent's active role in selecting and participating in environments that increase risk for poor outcomes. Although adolescence is a period of increased risk for substance use initiation, and experimental use of alcohol and tobacco can be considered normative, protective

factors (e.g. parental supervision, minimum legal drinking age) operate to keep overall levels of substance use low during adolescence. The variety of biological, social, and psychological changes that occur during adolescence highlight the importance of strategically timed, developmentally matched interventions that strengthen cognitive control over behavior in environments that minimize exposure to addictive substances.

TWO DEVELOPMENTAL PATHWAYS OF RISK FOR EARLY SUBSTANCE INVOLVEMENT

Two main pathways of risk for early substance involvement have been identified: *externalizing* (e.g. behavioral undercontrol, impulsivity, conduct problems) and *internalizing* (e.g. negative emotionality, depression and anxiety, sequelae of sexual and physical abuse) pathways. In one pathway, externalizing behaviors, which include behavioral disinhibition, impulsivity, and novelty seeking, robustly predict early onset and problematic use of substances in longitudinal research with youth. Because externalizing behaviors are typically present in childhood, before the onset of substance use, youth with high levels of externalizing behavior can be targeted for early intervention. The second, internalizing risk pathway has not been consistently linked to risk for substance use. The mixed findings across longitudinal studies for the internalizing risk pathway may be due to differences, for example, in the ages studied, and how internalizing symptoms were measured. Youth with high levels of both externalizing and internalizing symptoms are at greater risk for substance use than youth with symptoms in a single pathway, because the two pathways seem to have an additive effect on risk for substance involvement. The presence of high levels of externalizing and/or internalizing symptoms can escalate the progression of substance involvement and increase its severity and duration. In brief, externalizing and internalizing pathways of risk, signs of which are typically evident before the onset of substance use, suggest the important role of the dysregulation of behavior (e.g. high sensitivity to reward, low response inhibition) and affect (e.g. high mood lability, chronic depressed mood) as key risk factors for adolescent substance involvement.

PEAK PERIODS OF RISK FOR SUBSTANCE USE ONSET

There are two peak periods of risk for substance use initiation: during early adolescence, coincident with pubertal maturation, and during the transition to young adulthood, when many youth start to live independently, begin college, and enter the workforce. Around age 12, the prevalence of substance use begins to increase and continues to rise into young adulthood (mid-twenties) before leveling off. The early period of risk suggests the potential roles of early pubertal maturation, the normative adolescent increase in risk-taking behavior, and changes in the social environment (e.g. transition from grade school to high school) in the onset of substance use. The later period of risk suggests the possible roles of greater autonomy (e.g. driving a car, earning money), increasing influence of peers relative to family, and changes in the social environment (e.g. transition to independent living) on risk for substance use.

SEQUENCING OF SUBSTANCE USE INITIATION

The substances most commonly used by adolescents include alcohol, tobacco, and marijuana. The sequence of substance use initiation typically begins with the use of alcohol and tobacco before marijuana use; use of marijuana often precedes the use of other illicit drugs. Rates of other illicit drug use, such as cocaine and opiates, are generally low among youth. However, misuse of prescription medications, such as opiate-based painkillers and stimulants, is increasing, particularly among younger adolescents, who may have access to these medications at home. The fairly predictable sequencing of substance use initiation has led to the "gateway hypothesis," which proposes that use of a substance early in the sequence causes use of the next substance in the sequence. Support for the gateway hypothesis, however, is mixed, since environmental factors such as availability and access to certain substances influence the sequencing of substance use initiation.

PREVALENCE OF ADOLESCENT SUBSTANCE USE

Alcohol Use

Alcohol is the substance most often used by adolescents. More than 40% of youth have initiated alcohol use by age 16, and roughly 80% report some alcohol use by age 18. These numbers are cause for concern because alcohol use before ages 14–15 is associated with increased risk for developing an alcohol use disorder, and for using other substances. Youth who report alcohol use tend to engage in a particularly risky pattern of use known as heavy episodic ("binge") drinking (HED), in which five (for males)/four (for

females) or more standard drinks (standard drink = one can of beer, one shot of liquor, or one glass of wine) are consumed in a row. A national survey in 2008 found that 15% of 16-year-olds reported HED in the past 30 days, with the largest increase occurring between ages 17 (20%) and 18 (28%).

Tobacco and Illicit Drug Use

Similar to alcohol, the prevalence of tobacco use increases with age during adolescence. A national school-based survey found that 7% of 8th graders, 12% of 10th graders, and 20% of 12th graders reported cigarette use in the past month. For illicit drugs, the type of illicit drug most commonly used in the past month differs by age. Among 12- to 13-year-olds, a national survey found that 1.4% of youth reported misuse of prescription medication (e.g. pain relievers, tranquilizers, stimulants) and 0.9% reported marijuana use. In contrast, among 14- to 17-year-olds, marijuana was the illicit drug used most often in the past month (6% among 14- to 15-year-olds and 13% among 16- to 17-year-olds), followed by misuse of prescription medication (3% among 14–15 year olds and 5% among 16–17 year olds). The majority (67%) of individuals who try marijuana do so between the ages of 12 and 17. The difference between younger and older adolescents in the type of illicit drug most often used points to the importance of access to and availability of substances in the immediate environment as an important risk factor in adolescents' substance involvement.

Narrowing Gender Gap and Ethnic Differences in Substance Use Prevalence

Historically, adolescent males generally have had higher rates of substance use compared with females. However, the gender gap in rates of adolescent substance use has narrowed, with females approaching or surpassing rates of use reported by males at certain ages, and for certain substances. For example, 2008 national household survey data indicated little difference among 12- to 17-year-olds in rates of recent alcohol use (14.2% of males, 15.0% of females) and cigarette smoking (9.0% of males, 9.2% of females). However, males had slightly higher prevalence of marijuana use compared with females (7.3% vs. 6.0%), and females had higher prevalence of prescription drug misuse compared with males (3.3% vs. 2.5%).

With regard to ethnic differences in substance use, national school-based survey data indicate that among the three largest ethnic groups in the United States (i.e. whites, blacks, Hispanics), black students, compared with white students, had lower rates of use for most substances, including lower rates of any illicit drug and cigarette use. Hispanic students had rates of use that were generally lower than that of white students, but higher than that of black students. By 12th grade, however, Hispanic students had the highest rates of use of certain drugs (e.g. crack, crystal methamphetamine). The narrowing gender gap and ethnic differences in substance use suggest the potential utility of gender-specific, culturally sensitive interventions to reduce risk in targeted subgroups of youth.

CONTEXTS AND MOTIVES FOR ADOLESCENT SUBSTANCE USE

Societal constraints on substance use, such as a minimum legal drinking age of 21 and legal tobacco purchase age of 18, contribute to differences between adolescents and adults in primary contexts and reasons for use. Adolescent substance use typically occurs in groups of similar age peers (including siblings), and often in locations where adult supervision is minimal or absent (e.g. city parks, unsupervised home settings). Research on adolescents' "activity spaces" indicates that the use of specific substances (e.g. alcohol, tobacco, marijuana, prescription drug misuse) varies with location (e.g. someone's home, at school) and with the companions present. The illicit nature of youth substance use suggests that use is often opportunistic and that prevention efforts need to modify environments that promote substance use (e.g. through greater adult supervision and strict enforcement of underage alcohol and tobacco legislation).

Understanding a youth's motives for substance use can reveal processes underlying escalation and desistance from substance use. Research on adolescents' motives for substance use has focused largely on alcohol and tobacco. For alcohol, starting in early adolescence, three frequently reported reasons for drinking include enhancing positive feelings, facilitating social interaction, and coping with negative emotions. Importantly, adolescent HED is associated with enhancement and coping reasons for drinking, which may reflect externalizing and internalizing risk pathways, respectively. Among adolescents experimenting with cigarette use, curiosity and peer influence are the reasons for use most often reported. In contrast, among adolescent regular smokers, pleasure, relaxation, and difficulty quitting are the reasons most often reported for cigarette use. In sum, adolescents' reasons for substance use, particularly during early stages, include enhancing positive affect and experiencing positive drug effects, as well as peer influence. As adolescent substance involvement increases, motives for use increasingly tend to include coping reasons (e.g. alcohol) and difficulties in controlling use and craving (e.g. cigarette use).

TRAJECTORIES OF ADOLESCENT SUBSTANCE USE

Prevalence data provide a snapshot of substance use among adolescents at a specific time point. Longitudinal data, however, are needed to describe change over time in an individual's pattern of substance use. Trajectories capture heterogeneity in the course of substance use across individuals and provide information on individual differences in onset of substance use, rate of change in substance use, and developmentally specific turning points in substance use (e.g. escalation associated with the transition to high school). Across longitudinal community-based studies of youth, the most common trajectory types include stable low, chronic high, developmentally limited, and later onset increasing trajectories of substance use. In community samples, the most common or modal trajectory during adolescence for alcohol, cigarette, and marijuana use involves experimental to moderate use of alcohol and tobacco, and no use to experimental use of marijuana. The least common youth trajectory type involves stable, heavy use of a substance (e.g. <10% of youth report stable, heavy alcohol use).

Adolescent substance users, compared with their adult counterparts, are more likely to engage in a pattern of polysubstance use, typically involving alcohol, tobacco, and marijuana. Among youth who report use of more than one substance, trajectories for alcohol, cigarettes, and marijuana use tend to change together over time. However, onset of marijuana use generally follows onset of alcohol and tobacco. Despite generally parallel trajectories across the substances used by an individual, frequency of use may vary across substances, and drug substitution effects (i.e. use of one drug increases as use of another decreases) may occur. The general tendency for substance use trajectories to show parallel trajectories suggests common risk factors (e.g. trait disinhibition) and mechanisms (e.g. peer substance use) that underlie the similarity in overall trajectory of use across substances (e.g. alcohol and tobacco).

PREDICTORS AND OUTCOMES OF ADOLESCENT SUBSTANCE USE

Risk and Protective Factors

The primary systems of influence on risk for substance use include the individual (e.g. genetic liability, prenatal substance exposure, attitudes toward substance use, sensitivity to drug effects, reasons for use) family (e.g. parental substance use, family conflict), and extended social environment (e.g. peer substance use, transition to high school and college, school-level prevalence of substance use, neighborhood enforcement of minimum legal drinking age, media portrayals of substance use). Risk factors have been shown to predict initiation and escalation of use. Protective factors operate to modify risk for substance use (e.g. high parental monitoring, positive bond to school). Membership in stable, heavy substance use trajectories has been linked to risks such as family history of substance use problems, the adolescent's temperament (e.g. impulsivity) and co-occurring psychopathology (e.g. conduct problems, depression, response to trauma), and peer substance use. Youth in low substance use trajectories tend to have a low overall number of risk factors, as well as the presence of protective factors. The adolescent shift in the importance of peers in relation to family suggests the critical role of peer selection and influence on adolescent substance use trajectories.

Outcomes Associated with Substance Use Trajectories

As may be expected, youth with trajectories representing no to low substance use into young adulthood tend to have the best outcomes in multiple areas of functioning, such as academic, interpersonal (family and peer), and physical and mental health domains. By contrast, youth with more chronic and severe trajectories of substance use, including increasing substance use during the transition to adulthood, tend to have poor social, academic, employment, legal, and health outcomes in adolescence (e.g. greater risk for sexually transmitted disease, substance-related injury and auto accidents). These adverse outcomes may persist into young adulthood, and compromise young adult functioning. In addition, there is emerging evidence that suggests the negative effect of alcohol consumption on the developing human brain. For example, HED among adolescents is associated with damage to white matter tracts that connect regions in the hippocampus and prefrontal cortex. In addition, youth who report HED, compared with nondrinking peers, showed subtle cognitive impairments on a verbal learning task. Although these findings are suggestive of alcohol's effects on the developing brain, further research is needed to determine the persistence of these effects, and the extent to which such differences existed before or were exacerbated by substance use.

DIAGNOSTIC AND STATISTICAL MANUAL DEFINITION OF SUD

Core features of addiction are proposed to involve physical, behavioral, and psychological components.

These core features are distinguished from ancillary problems (e.g. interpersonal or legal problems) that can occur as a consequence of substance use. The *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition* (DSM-IV), recognizes two SUDs: abuse and dependence. The two SUDs attempt to preserve the distinction between substance-related negative consequences of use (i.e. abuse diagnosis), and a pathological pattern of compulsive drug seeking and use that indicates addiction (i.e. dependence diagnosis). DSM-IV substance abuse is defined by certain negative consequences resulting from substance use (e.g. substance-related interpersonal problems; impaired functioning at school, home, or work due to substance use; substance-related legal problems) or hazardous substance use (e.g. repeated episodes of driving while intoxicated). The presence of one of four criteria is needed for an abuse diagnosis. The DSM-IV substance-dependence diagnosis is intended to identify a compulsive pattern of substance use (i.e. addiction), in which three of seven criteria are present within the same 12-month period. Dependence criteria represent physical symptoms (i.e. tolerance or withdrawal), salience of drug-taking behavior (e.g. much time spent using, reduced activities to use), and impaired control over substance use (e.g. using more or longer than intended, difficulties quitting or cutting down on use, continued use despite physical or psychological problems caused by substance use). The diagnosis of a DSM-IV SUD requires evidence of clinically significant impairment or subjective distress resulting from substance use. The substance abuse diagnosis is generally considered a milder illness relative to dependence, due to its one symptom threshold, and because a diagnosis of dependence precludes a diagnosis of abuse, indicating the hierarchical relation between the two disorders.

LIMITATIONS OF DSM-IV SUDS

Valid diagnosis provides a standard operational definition of an illness that provides information on cause, prognosis, and treatment. DSM-IV–defined SUDs represent evolving working definitions, which were developed based primarily on clinical observation and empirical findings involving adult patients. Research generally supports the reliability and validity of DSM-IV abuse and dependence diagnoses. However, important limitations of these diagnoses have been identified. First, diagnoses impose relatively arbitrary distinctions (i.e. illness present or absent) on a continuum of substance use severity. The thresholds are artificial constructions, and result in symptomatic cases that nevertheless fall short of the threshold for illness. Second, because no symptom is necessary for a diagnosis, a high degree of heterogeneity exists among those with the same SUD diagnosis, which complicates efforts to identify the cause and improve treatment. Third, DSM-IV criteria do not correspond well to emerging research findings on the neurobiology of addiction, which suggests important roles for reward processing and motivation, deficits in cognitive functioning (e.g. response inhibition, memory), and dysregulation of behavior and affect. Although there is some consensus regarding a basic definition of addiction, fundamental questions remain regarding how to optimally define and assess its core features and boundary conditions.

Limitations of DSM-IV SUDs that are more specific to youth include thresholds that were developed for use with adults, and the need to consider symptoms and symptom manifestations that are relevant to youth substance involvement. The application of DSM-IV SUD thresholds to identify youth with an SUD may result in under-identification of problem use in youth, because lower levels of problem severity, compared with adults, may indicate treatment need among adolescents. In addition, the type of symptoms used to identify DSM-IV SUDs do not cover the types of problems most likely to manifest in young substance users, particularly during early stages of problem use. For example, youth who are beginning to experience substance-related problems are less likely to endorse a strong desire or attempts to limit use. Further, some symptoms, such as substance-related legal problems, may have different meaning in adolescents, for whom alcohol use under age 21 is a status offense. Other symptoms, such as driving when intoxicated as an indicator of hazardous use, also are less applicable to youth who do not drive. These limitations of DSM-IV SUDs can result in under-identification of youth who may benefit from intervention.

DEVELOPMENTALLY TAILORED ASSESSMENT OF SUBSTANCE INVOLVEMENT IN YOUTH

Adolescent substance users, compared with adults, tend to have shorter histories of use and generally experience milder substance-related problems. These developmental differences emphasize the need to adapt addiction constructs and SUD criteria to make them relevant to, and properly scaled for, an adolescent's developmental stage. Specifically, certain symptoms can manifest and be interpreted differently in adolescents and adults. For example, youth may report the dependence symptom "spending much time trying to obtain alcohol, drinking, or getting over its effects" due to difficulties in obtaining alcohol due to minor status, rather than as an indicator of a compulsive pattern of use. Other symptoms need to be appropriately scaled to

indicate a problematic level of use in youth. For example, a high level of "tolerance" indicating dependence is challenging to assess in adolescents, because on-going physical development, rather than addiction-related processes, may explain increases in consumption to obtain the same effect. Alternatively, some youth may report dependence symptoms (e.g. withdrawal) at low and infrequent levels of use, particularly in relation to tobacco. In these situations, it is unclear whether inexperience with the substance and the adolescent's interpretation of the symptom results in a "false-positive" endorsement of the symptom, or whether the youth's self-report indicates an early stage of addiction, albeit at a very low level of use that is not traditionally considered to indicate addiction. Valid assessment of substance-related problems in youth involves tailoring symptom assessment to the adolescent's developmental and social context, and appropriately scaling symptoms to minimize false-positive and false-negative symptom and diagnostic assignments.

DEVELOPMENT OF SUDS IN ADOLESCENTS

Among adolescents who initiate use and progress to SUD, time to SUD onset occurs, on average, within roughly 3 years of initiation. Despite individual differences in rate of symptom development, and type of symptoms experienced, there is some regularity in the sequential emergence of SUD symptoms in adolescents. For alcohol, alcohol-related interpersonal problems tend to emerge within the first 2 years of the onset of regular drinking, followed by other alcohol-related negative consequences and symptoms of dependence, with withdrawal typically emerging last. For tobacco, the first symptoms tend to emerge within 1 year of the onset of monthly use and typically involve symptoms of impaired control over cigarette use (e.g. strong craving). For marijuana, similar to tobacco, within a year of the onset of regular use, a symptom of impaired control over substance use (i.e. use more than intended) tends to emerge first, followed by the onset of marijuana-related physical or psychological problems 1 year later. Cross-drug differences in order of symptom onset suggest possible differences in a substance's addictive liability, such that early report of impaired control over use for tobacco and marijuana could be interpreted to indicate their higher addictive liability, relative to alcohol, among youth. However, the extent to which early self-reports of impaired control over substance use truly indicate a compulsive pattern of use, particularly during relatively early stages of use, or difficulties in controlling use that may be related to trait impulsivity (rather than compulsion), remains to be determined. Findings regarding symptom development need to be interpreted with caution given the need to consider methodological issues in symptom assessment among youth.

PREVALENCE OF ADOLESCENT SUDS

Similar to the prevalence of substance use, SUD prevalence increases with age during adolescence, and peaks in young adulthood. The most prevalent SUDs among youth involve alcohol, tobacco, and marijuana. In 2008, national survey data indicated that 7.6% of 12- to 17-year-olds met DSM-IV SUD criteria in the past year (most often for the milder abuse diagnosis), compared with 20.8% of 18- to 25-year-olds. Up to an additional 17% of youth are estimated to be "diagnostic orphans," who report substance-related problems but do not meet criteria for an SUD. Orphans may have up to two DSM-IV dependence criteria, and although they generally report a level of substance use similar to youth who meet criteria for abuse, orphans do not qualify for an SUD diagnosis. Orphans are of particular interest because the DSM-IV dependence diagnosis represents a "mature" or fully developed phenotype, typically observed in adult substance users, after many years of heavy use. Youth who are sub-threshold for SUD (e.g. orphans) are at risk for progression to SUD and warrant intervention and follow-up to halt progression to full-blown SUD. Certain settings (e.g. juvenile justice, psychiatric) report high rates of adolescent SUD, relative to the general population. Compared with adults, adolescent substance users are more likely to be in early stages of problem use and to report milder substance-related problems that highlight the need for early identification and intervention to reduce risk for progression and adverse outcomes.

ADOLESCENT SUD COURSE IN COMMUNITY SAMPLES

Longitudinal epidemiologic studies indicate that risk for more chronic alcohol, tobacco, and marijuana use disorders is concentrated within the subgroup of individuals with early onset of substance use (e.g. before age 14–15 for alcohol, age 16 for marijuana). These early onset youth are more likely to establish a regular pattern of substance use during adolescence, which has been associated with greater risk for progression to chronic and heavy substance use. Furthermore, studies of adolescent substance use trajectories indicate that among early-onset substance users, although some show developmentally limited patterns of substance use, others show a stable high level of use and associated risk for SUD into adulthood. The persistence and

stability of heavy substance use during adolescence is robustly predicted by temperament (i.e. behavioral disinhibition) and, relatedly, co-occurring psychopathology (e.g. conduct problems).

ADOLESCENT SUD TREATMENT

Among youth with an SUD, there is high unmet treatment need. Most (~90%) youth with an SUD in a 2008 national survey reported receiving no treatment. In part, low rates of adolescent treatment utilization reflect the fact that most youth do not refer themselves to treatment, and are typically mandated to treatment (e.g. by court, school), or attempt to reduce substance use on their own. Although many youth with lower levels of substance use severity are able to reduce use on their own, those who engage in heavy substance use warrant formal intervention.

Treatment services for substance using youth include a continuum of care ranging from brief interventions (e.g. 20-min feedback to motivate efforts to stop substance use), outpatient, intensive outpatient, partial hospitalization, inpatient, and residential placement. Most youth who receive treatment for substance use are treated at the outpatient level of care. Substance use treatment for youth typically involves the adolescent's family and often includes consultation with psychiatric and medical specialists to address the adolescent's problems in multiple areas of functioning (e.g. depression, sexually transmitted disease). More than half of the youth in substance use treatment are estimated to have a co-occurring psychiatric condition (e.g. conduct problems, depression, trauma-related), which may predate, result from, or be exacerbated by heavy substance use. Among empirically supported treatments, no single type of treatment content for youth (e.g. 12-step based cognitive-behavioral) has been shown to be clearly superior to any other. However, receipt of treatment results in better outcomes compared with wait-list control groups, and both longer duration of treatment and treatment completion predict better outcomes. Due to a history of heavy substance use, some youth may require more than one episode of treatment, as well as participation in a program of continuing care, to stably maintain long-term reductions in substance use.

POST-TREATMENT TRAJECTORIES OF ADOLESCENT SUBSTANCE INVOLVEMENT

Post-treatment Substance Use Trajectories

Although treated adolescents generally show reductions in substance use, there is variability in the course of adolescents' post-treatment substance use. Research on post-treatment trajectories of adolescent substance involvement has identified stable low and stable high levels of use, and decreasing, as well as increasing (for a minority of youth) patterns of use for alcohol and marijuana. In the year after substance use treatment, many youth (roughly half) are in abstinent or infrequent use trajectories, with smaller proportions classified in persistent high or slowly decreasing substance involvement trajectories. Analyses of concurrent trajectories of change after treatment for alcohol, marijuana, and other drugs indicate a moderate level of cross-drug concordance in a pattern of change over a 3-year follow-up. The similarity in pattern of change across substances (excluding tobacco, which tends to show little change after substance use treatment) provides support for similar underlying mechanisms of change after treatment, for most substances.

Outcomes Associated with Post-treatment Trajectories

As may be expected, youth in stable post-treatment abstinence and low use trajectories had better interpersonal, health, and academic/work outcomes compared with youth in chronic heavy use trajectories. More chronic trajectories frequently reflect the adverse effect of co-occurring psychopathology in the course of adolescent substance use. Changes occurred at different rates for various areas of functioning, with improvement in school performance occurring within a year after treatment, but improvement in family functioning was only evident 2 years after treatment. Despite reductions in substance use, treated youth continued to show greater problem severity in a number of domains, compared with a community comparison sample, suggesting the adverse effect of adolescent substance use.

Predictors of Post-treatment SUD Course

During treatment, factors such as family involvement and higher motivation to abstain predict better adolescent substance use outcomes. Because most youth are mandated to treatment, youth may be reluctant to reduce substance use and to "give up" friends and a "lifestyle" that supported substance use. Thus, an important role for treatment involves boosting and supporting an adolescent's motivation to reduce substance use and helping youth to identify alternative healthy social activities and environments. Post-treatment factors, such as affiliation with substance using peers and ability to cope with stressors, account for more of

the variance in long-term treatment outcome than before and during treatment factors.

Studies of reasons for relapse indicate differences between adolescents and adults in the reasons most often reported as a precipitant of relapse. Among youth, social situations (e.g. attending a party where alcohol and marijuana were present) and peer use are frequently reported relapse reasons for alcohol and marijuana. In contrast, among adults, craving and relief of negative emotions are often cited as reasons for relapse. Differences in relapse reasons suggest the relative importance of the social environment in influencing relapse for youth, in accord with youths' generally milder substance use severity, whereas coping and compulsion-based reasons predominate for adults. Among youth attempting to stop cigarette use, however, compulsion-based reasons for relapse are often reported. These findings suggest the generally milder severity of substance involvement in youth, compared with adults, and among youth, suggest cross-drug differences as reasons for relapse that seem to be linked to the degree of addiction severity.

IMPLICATIONS FOR IMPROVING ADOLESCENT SUBSTANCE USE PREVENTION AND INTERVENTION

The approach to addiction as a developmental disorder highlights the importance of prevention and early intervention among youth to reduce substance-related harm. Developmental changes that are specific to adolescence result in a convergence of risk for initiation of substance use. The two main pathways of risk represented by externalizing and internalizing symptoms, suggest targeting youth with high levels of these symptoms for early intervention to delay substance use onset and halt progression of use. In addition to targeted intervention, universal prevention efforts that are coordinated across national and community levels are needed to strengthen protective factors (e.g. enforcement of drug control policy, effective parenting) that aim to minimize environmental risks for adolescent substance use. Although environmental factors (e.g. drug availability) play an important role in risk for substance use onset, as substance use progresses, individual factors (e.g. externalizing behavior, reasons for use), in combination with the adolescent's social environment, become increasingly important targets for intervention. Adolescent substance use trajectory data can usefully inform the timing of interventions. Emerging findings on adolescent brain development, reward sensitivity and decision-making processes, and social cognition can be used to develop interventions specifically tailored to maturational stage and salient social contexts. Research is needed to better specify the neurobiological pathways and environmental contexts involved in the onset, maintenance, and desistance from addiction, specifically in relation to developmental stage of maturation and the shift to compulsive drug taking.

SEE ALSO

Symptoms and Course: Alcohol Use Disorder in Adulthood, Developmental Risk Taking and the Natural History of Alcohol and Drug Use among Youth, Epidemiology of Adolescent and Young Adult Alcohol, Tobacco, and Drug Use and Misuse in the United States

Glossary

Addiction a compulsive pattern of substance use, often maintained despite significant personal harm (e.g. physical or psychological problems due to use) or impairment in functioning caused by substance use.

Diagnostic and Statistical Manual (DSM) of Mental Disorders the DSM is published by the American Psychiatric Association and provides operational definitions for common mental disorders. DSM is currently in its fourth edition (DSM-IV; 2000); a fifth edition will be published in 2013. DSM-IV recognizes two substance use disorders: abuse and dependence (see below).

Diagnostic orphans individuals who are symptomatic (e.g. may have up to two DSM-IV dependence criteria for a given substance), but who do not qualify for a substance use disorder diagnosis. Diagnostic orphans illustrate a limitation of the DSM-IV classification for substance use disorders.

Externalizing behaviors behaviors that reflect impulsivity or behavioral under-control, high novelty or sensation seeking, conduct problems and aggression, and delinquency.

Heavy episodic ("binge") drinking (HED) episodes of high-volume alcohol consumption in which at least five (for males) or four (for females) standard drinks (see below) are consumed in a row.

Internalizing behaviors behaviors that reflect negative mood states such as depression and anxiety, and/or an interpersonal style that could be described as withdrawn and inhibited.

Sensitization increased sensitivity to drug effects with repeated dosing, which may be long lasting (e.g. substance-induced changes in gene expression).

Standard drink standard unit of measure for alcohol consumption that is equivalent to one 12-ounce can of beer, one shot (1.5 ounces) of 80-proof liquor, or one 5-ounce glass of wine (12–14% alcohol).

Substance abuse one of the two DSM-IV substance use disorders, which involves meeting one of the four criteria related to certain negative consequences of substance use (i.e. impairment in role functioning due to substance use, hazardous substance use, substance-related legal problems, substance-related interpersonal problems).

Substance dependence one of the two DSM-IV substance use disorders, which requires that three or more of seven criteria (i.e. tolerance, withdrawal, using more or longer than intended, strong desire or failed attempts to limit use, much time spent using, reduce activities to use, substance-related psychological or physical problems) co-occur within a 1-year period.

SUDs substance use disorders.

Acknowledgments

Preparation of this work was supported by National Institute on Alcohol Abuse and Alcoholism K02 AA018195, R01 AA014357, R21 AA017128, and National Institute on Drug Abuse R01 DA012237.

Further Reading

Brown, S.A., McGue, M., Maggs, J., et al., 2008. A developmental perspective on alcohol and youths 16 to 20 years of age. Pediatrics 121, S290–S310.

Chung, T., Maisto, S.A., 2006. Relapse to alcohol and other drug use in treated adolescents: review and reconsideration of relapse as a change point in clinical course. Clinical Psychology Review 26, 149–161.

Clark, D.B., Thatcher, D.L., Tapert, S.F., 2008. Alcohol, psychological dysregulation, and adolescent brain development. Alcoholism: Clinical and Experimental-Research 32, 375–385.

Maggs, J.L., Schulenberg, J.E., 2005. Initiation and course of alcohol consumption among adolescents and young adults. In: Galanter, M. (Ed.), Recent Developments in Alcoholism. Alcohol Problems in Adolescents and Young Adults, vol. 17. Kluwer Academic/Plenum Publishers, New York, NY, pp. 29–47.

Spoth, R., Greenberg, M., Turrisi, R., 2008. Preventive interventions addressing underage drinking: state of the evidence and steps toward public health impact. Pediatrics 121, S311–S336.

Waldron, H., Turner, C., 2008. Evidence-based psychosocial treatments for adolescent substance abuse: a review and meta-analysis. Journal of Clinical Child and Adolescent Psychology 37, 1–24.

Windle, M., Spear, L.P., Fuligni, A.J., et al., 2008. Transitions into underage and problem drinking: developmental processes and mechanisms between 10 and 15 years of age. Pediatrics 121, S273–S289.

Zucker, R.A., 2006. Alcohol use and the alcohol use disorders: a developmental biopsychosocial systems formulation covering the life course. In: Cicchetti, D., Cohen, D.J. (Eds.), Developmental Psychopathology. Risk, Disorder, and Adaptation, vol. 3. Wiley, Hoboken, NJ, pp. 620–656.

Relevant Websites

http://www.oas.samhsa.gov/nhsda.htm – National Household Survey on Drug Use and Health.

http://monitoringthefuture.org/ – Monitoring the Future.

http://www.niaaa.nih.gov/Resources/RelatedWebsites/ – National Institute on Alcohol Abuse and Alcoholism.

http://www.nida.nih.gov/nidahome.html – National Institute on Drug Abuse.

http://www.drugabuse.gov/PODAT/PODATIndex.html – Principles of Drug Addiction Treatment: A Research-Based Guide.

http://www.iom.edu/Reports/2009/Preventing-Mental-Emotional-and-Behavioral-Disorders-Among-Young-People-Progress-and-Possibilities.aspx – National Academy of Sciences: Preventing Mental, Emotional, and Behavioral Disorders among Young People: Progress and Possibilities.

http://www.aacap.org/galleries/PracticeParameters/JAACAP%20Substance%20use%202005.pdf – American Academy of Child and Adolescent Psychiatry (AACAP): Practice Parameter for the Assessment and Treatment of Children and Adolescents with Substance Use Disorders.

CHAPTER

11

Symptoms and Course: Alcohol Use Disorder in Adulthood

*Michie N. Hesselbrock**, *Karen Grube Chartier*[$], *Grace Chan**, *Victor M. Hesselbrock**

[*]University of Connecticut School of Medicine, Farmington, CT, USA

[$]University of Texas, Dallas, TX, USA

OUTLINE

INTRODUCTION

Alcohol is the most widely used drug in the world. Beverage ethanol is legally sold in the United States, and despite age restrictions on sales to those less than 21 years old many begin drinking at an early age, often before age 15. Alcohol is used in a variety of ways and the patterns and amounts of drinking vary among drinkers (e.g. infrequent, regular, binge, and heavy drinking). The ingestion of alcohol produces a variety of actions on the brain and behavior. The effects can be acute, caused immediately after the ingestion of alcohol, or chronic, resulting from repetitive heavy drinking. The acute effects of alcohol are dose dependent, with higher doses of alcohol producing larger effects on biological process and a more marked effect on behavior than lower doses of alcohol. While the acute effects of alcohol can be temporary, the continuation of regular intoxication can lead to chronic heavy drinking and alcohol dependence. Alcohol dependence is associated with adverse physical and psychological health, as well as family and other social consequences. This chapter will review the symptoms and conditions associated with heavy drinking and alcohol dependence and their effects on the course of alcohol use disorder.

HEAVY DRINKING AND ALCOHOL USE DISORDERS

The pattern of consumption most notably associated with chronic alcohol-related problems is heavy drinking. Heavy alcohol use, as defined by the National Institute on Alcohol Abuse and Alcoholism, includes consuming five or more standard drinks per day (or 15 or more per week) for men and four or more drinks

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00011-5

per day (or eight or more per week) for women. One standard drink is equivalent to 12 ounces of beer, 5 ounces of wine, or 1.5 ounces of 80-proof spirits. Frequent heavy drinkers tend to develop alcohol dependence.

According to the DSM-IV-TR that was published by American Psychiatric Association (APA), alcohol dependence is characterized by a maladaptive pattern of alcohol use, leading to clinically significant impairment or distress. The DSM-IV criteria for alcohol dependence require at least three of seven symptoms to be clustered in a 1-year period. These symptoms include two indicators of physical dependence (i.e. tolerance and withdrawal), three indicators of loss of control (e.g. drinking alcohol more than intended), a persistent desire or unsuccessful efforts to cut down or control use, and continued use despite physical or psychological problems caused by drinking. There are two additional indicators for spending excessive time on drinking-related tasks: A great deal of time spent on activities necessary to obtain, use, and recover from the effects of alcohol and giving up important activities due to drinking. The DSM-IV criteria specify early and sustained full remission from alcohol dependence, when criteria for dependence are no longer met. Early remission is defined as no positive symptoms for 1–12 months, while sustained remission lasts for a period of 12 months or longer. Partial remission is specified when one or more symptoms are present, but do not meet dependence or abuse criteria. The severity of alcohol dependence and alcohol-related conditions vary from person to person, depending on gender, age, family history of alcohol dependence, comorbid psychiatric conditions, and the amount and duration of heavy drinking.

The proposed DSM-V criteria, to be published by APA in 2013, will make several changes to the DSM-IV/DSM-IV-TR. First, both alcohol dependence and alcohol abuse will be subsumed under the single label of alcohol use disorder. Currently under the DSM-IV, alcohol abuse is a separate diagnosis from alcohol dependence specified by one of four symptoms for abuse (i.e. recurrent alcohol use resulting in a failure to fulfill major role obligations, recurrent use in situations in which it is physically hazardous, recurrent substance-related legal problems, continued substance use despite having persistent social or interpersonal problems caused or exacerbated by the effects of alcohol). The DSM-V will use 11 criterion symptoms to define an alcohol use disorder; the current legal problems symptom for alcohol abuse will be excluded from the DSM-V alcohol use disorder criteria and a new craving symptom will be added. An alcohol use disorder is characterized by "maladaptive patterns of alcohol use leading to clinically significant impairment or distress as manifested by 2 of 11 symptoms within a 12-month period." The DSM-V will also use new specifications for severity; a moderate diagnosis is 2–3 criteria positive and a severe diagnosis is 4 or more criteria positive. The new criteria will specify an alcohol use disorder with and without physiological dependence by the presence or absence of tolerance or withdrawal. The APA is currently conducting a field trial and soliciting feedback from the community on these changes.

EARLY ONSET OF DRINKING AND THE DEVELOPMENT OF ALCOHOL DEPENDENCE

There is an abundant empirical literature supporting an early age of onset of drinking as a major risk factor for the development of alcohol dependence, from both retrospective and longitudinal studies. Young and colleagues surveyed 3072 adolescents drawn from three community-based family samples in Colorado. According to the survey, by 18 years of age, 1 in 10 adolescents abuse alcohol and 3.5% of adolescents, age 12–18 years old meet criteria for alcohol dependence. The National Survey of Drug Use and Health (NSDUH) in 2004 found that among 14 million alcohol dependent adults aged 21 or older, more than 13 million started using alcohol before age 21. Further, the NSDUH showed that early onset alcohol users (first use before age 15) were five times more likely than later onset alcohol users (after age 21) to develop alcohol dependence. A 16-year prospective follow-up study of adolescents also reported that both moderate and drunkenness-oriented drinking at 16 years of age are associated with excessive alcohol use in adulthood.

An early age of onset of drinking also affects the developmental course of alcohol dependence. The epidemiological literature shows that many adolescents who experiment with alcohol at an early age also develop alcohol dependence and related problems at an early age. Data from the 2001–2002 National Epidemiological Survey of Alcohol Related Conditions (NESARC) show that an early drinking onset is associated with an onset of alcohol dependence within 10 years of the first drinking experience. Consequently, those who develop alcohol dependence at an early age experience severe alcohol dependence symptoms and related health and psychiatric conditions. Hingson and colleagues reported that individuals who develop alcohol dependence before age 18 experience chronic relapsing, multiple, and longer dependence episodes with a wider variety of symptoms in the NESARC data.

The association between an early age of onset of alcohol problems and the severity of symptoms has also been documented in clinical samples. Bucholz and

her colleagues conducted a latent class analysis of 36 alcohol dependence symptoms in a sample of biological relatives of alcohol-dependent probands in the Collaborative Studies on Genetics of Alcoholism (COGA). This study found that the cluster characterized by an early age of onset of regular drinking and/or intoxication was associated with an early onset of alcohol problems and more severe alcohol-related symptoms.

Further, the men who began having alcohol problems before age 25 as compared to those who experienced problems at a later age, tended to use other drugs and experienced a greater number of drug-related problems. Similarly, Schuckit and colleagues found that alcohol-dependent persons whose alcohol diagnosis was preceded by antisocial personality disorder tended to begin to have alcohol problems before age 25. Despite their younger age, they had heavier alcohol involvement, drug abuse, and social and legal problems and had contact with treatment at a younger age.

In summary, there is consistent evidence linking an early age of onset of alcohol use and the development of alcohol dependence at an early age. In addition, an early age of onset of alcohol use is also associated with the severity and a relapsing course of alcohol dependence.

GENDER AND THE PROGRESSION OF ALCOHOL DEPENDENCE

Men compared to women are generally more likely to be alcohol dependent and to have a more severe course of alcohol dependence. For example, in the COGA sample, men report an earlier onset of weekly drinking and more alcohol use disorder episodes. However, drinking-related problems appear to develop more rapidly for women than men in another study. Studies support a "telescoping" effect in the progress of alcohol dependence among women, characterized by a shorter time from the onset of drinking to treatment entry and by the earlier onset of alcohol-related complications. The study of alcohol-dependent women with opioid and cannabis dependence also found that women had a faster progression to treatment entry than men despite no gender difference in the age at onset of regular use of any substance. Further, women reported greater psychiatric and employment problems and comparable alcohol and drug severity and medical problems compared to men at treatment entry, despite showing fewer pretreatment years of use. Similar findings were reported in COGA subjects. Despite reporting a later age of regular drinking and onset of alcohol dependence, women sought help from a health professional and reported the physiological symptoms of shakes and full withdrawal at a similar age to men. The average delay between the onset of dependence and initiation of help was 4 years for women and 6 years for men. However, despite these differences alcoholic men and women had many similarities in the course of their disorders. The progression of symptoms, or order in which problems occurred did not differ dramatically by gender.

The literature thus far indicates that while there is evidence of a telescoping effect in the development of alcohol dependence and time to first treatment among women, there is little evidence of gender differences in terms of the natural history of alcohol dependence.

ALCOHOL DEPENDENCE AND PHYSICAL HEALTH

There is well-established evidence regarding the direct causal associations between chronic heavy drinking and excess morbidity and mortality. Many studies document the effect of chronic heavy alcohol use on a wide range of adverse physical conditions, including acute hepatitis, alcoholic cirrhosis, diabetes, degeneration and atrophy of nerve cells, stroke, certain malignancies, traumatic injuries, cardiomyopathy, hypertension, gastritis, pancreatitis, muscle weakness and pain, infertility, and gestational problems. Diseases related to the direct toxic effects of ethanol use include the liver conditions of acute hepatitis and alcoholic cirrhosis, cardiomyopathy, degenerative disease of the heart muscle, injury to skeletal muscles, and diseases of the gastrointestinal tract including gastritis and pancreatitis. Other chronic effects can be attributed to specific vitamin deficiencies (e.g. thiamine). Thiamine deficiency, common in chronic alcoholics, has been linked to such nervous system conditions as degeneration of nerve cells, disturbed cognition, and severe memory loss (including Korsakoff's).

While moderate alcohol consumption has been consistently associated with reduced risk for cardiovascular diseases by raising high-density lipoprotein cholesterol and reducing plaque accumulations in the arteries, chronic heavy drinking can lead to high blood pressure, alcoholic cardiomyopathy, congestive heart failure, and hemorrhagic stroke. Heavy drinking also raises triglyceride levels. Significant increased risks were found at higher rates of alcohol intake for coronary heart disease and ischemic stroke at 100 g day^{-1} and for hemorrhagic stroke at 50 g day^{-1}. Excessive alcohol consumption is 1.5–2 times more likely to affect high blood pressure. This association increases dramatically when alcohol intake exceeds 5–6 drinks per day.

Chronic heavy drinking is also associated with an elevated risk for various types of cancer. It is estimated that 2–5% of all cancer cases are directly or indirectly related to alcohol use. An age-specific distribution of

adult drinkers obtained from the World Health Organization's (WHO) Global Burden of Disease project was used to estimate the number of cancer cases deaths linked to alcohol use. Buffeta and colleagues attributed 3.6% of all cancers (5.2% in men and 1.7% in women) and 3.5% of all cancer deaths to heavy drinking. Among women, breast cancer accounted for 60% of alcohol-attributed cancers, while for men 60% of cases were cancers of the oral cavity, pharynx, and esophagus. A meta-analysis of 156 studies involving 116 702 subjects also found strong direct risks for cancers of the oral cavity, esophagus, and larynx, as well as hypertension, liver cirrhosis, chronic pancreatitis, and injuries due to heavy alcohol consumption. Direct relationships with heavy drinking were also found for cancers of the colon, rectum, liver, and breast, although these risk trends were not as strong. Significant increased risk for all conditions were found beginning at 25 g day^{-1} of alcohol intake, corresponding to approximately two drinks per day.

Empirical studies thus far indicate several consistent findings. First, moderate alcohol consumption is associated with a reduction in cardiovascular disease and mortality in the general population. A second consistent finding is the wide variety of adverse health effects that result from chronic heavy drinking. These findings are consistent across different countries.

ALCOHOL DEPENDENCE AND PSYCHIATRIC COMORBIDITY

Chronic heavy alcohol consumption and alcohol dependence are not only associated with poor physical health, but also with a variety of psychiatric conditions. Findings from the general population 2001–2002 NESARC indicated that among respondents with a current alcohol use disorder, 13.7% had major depression, 17.1% had any anxiety disorder, and 28.6% had at least one personality disorder, with antisocial personality disorder (ASPD) being the most common. Anxiety and mood disorders were more prevalent in respondents diagnosed with current alcohol dependence (any anxiety, 23.25%; any mood, 27.55%) as compared to those diagnosed with alcohol abuse (any anxiety, 11.81%; any mood, 11.73%).

Psychiatric comorbidity has also been found among clinical samples. Major depressive disorder was the most prevalent comorbid disorder reported, with lifetime diagnosis rates ranging from 52 to 70% among women and 34 to 55% among men who received inpatient treatment for alcohol dependence. The majority of depressive symptoms were alcohol related. The rate of ASPD was also high, ranging from 25 to 39% for men and 11 to 31% for women in the clinical sample.

There are gender differences in comorbid psychiatric disorders. Women with alcohol use disorders in both treatment and community samples compared to men have higher rates of axis I co-occurring mental disorders. Watkins studied patients in outpatient treatment and found that women are more likely than men to have co-occurring mental health and substance use disorders. Depressive and anxiety disorders were more common in women with alcohol use disorders, while ASPD was more prevalent in men. In the 2001–2002 NESARC study, 10% of men as compared to 7% of women with lifetime alcohol use disorders met DSM-IV criteria for ASPD. However, despite men showing a higher prevalence of ASPD than women, men and women with co-occurring ASPD and alcohol use disorders reported similar ages at first drink and at the onset of an alcohol use disorder, and similarities in the duration of longest alcohol use disorder episode and the volume and patterns of alcohol consumption during the period of heaviest drinking.

Comorbid mental health and alcohol use disorders are associated with more severe course of alcohol dependence and greater alcohol-related problems. Depression and ASPD, in particular, have been recognized to negatively impact alcohol use and related conditions among alcohol-dependent persons. In the 2001–2002 NESARC study, alcohol use was most severe for individuals with co-occurring lifetime alcohol use disorders and ASPD. Individuals with ASPD reported a greater prevalence of family history of alcohol dependence (82.1%) than individuals with no antisocial behavior syndrome (63.0%), as well as younger ages of first drink (mean: 15.3 versus 17.9 years), onset of alcohol use disorders (mean: 19.4 versus 23.3 years), and first treatment episode (mean: 25.9 versus 33.7 years). Compared to the no ASPD group, individuals with ASPD were more likely to have two or more episodes of alcohol use disorder (45.9 versus 28.8%), and to drink both more (mean ounces of ethanol/drinking day: 6.79 versus 3.14 oz) and more often (>312 days $year^{-1}$: 35.8 versus 19.3%) during the period of heaviest drinking.

While depressive disorders are associated with the severity of alcohol dependence, the COGA study found little evidence that depressive symptoms lead to the initiation of alcohol use among adolescents. However, once heavy regular alcohol use has begun, depressive symptoms and major depressive disorder predict an early age of onset of alcohol dependence and the severity of drinking problems. Many studies show strong associations between depression and relapse to alcohol use after treatment. Depression and the severity of depressive symptoms were associated with an increased risk for relapse among Veterans Administration patients. Greenfield and colleagues found that subjects diagnosed with comorbid depression relapsed

more quickly after treatment. A study conducted in Iceland similarly showed that the co-occurring diagnoses of alcohol dependence and anxiety disorder were associated with relapse after the treatment; patients with no comorbid psychiatric disorders had fewer lifetime detoxification center admissions. Agoraphobia and panic disorder predicted readmission to a detoxification center among treatment recipients with fewer than two prior admissions, while polysubstance abuse was associated with three or more readmissions.

An association between relapse to depression and drinking has been supported by long-term follow-up studies. Hasin and colleagues analyzed 5-year follow-up data among patients participating in the Collaborative Study of Depression who were diagnosed with comorbid alcohol dependence. They identified a persistent depressive disorder as a significant predictor of relapse to alcohol use, while improvement in alcohol use was associated with reduced depressive symptoms.

Depressive symptoms that are alcohol induced should be distinguished from symptoms independent (not associated) of alcohol use among alcohol-dependent persons, as they may be differentially associated with the course of alcohol dependence. In a COGA sample, alcohol-induced depression was associated with persistent alcohol dependence and alcohol-induced depression 5 years later, while nonalcohol-induced depression was associated with persistent nonalcohol-induced depression.

The review of the literature clearly shows a high prevalence of psychiatric comorbidity and alcohol dependence. It is also clear that psychiatric comorbidity has significant adverse effects on the course of alcohol dependence with increased risk for relapse.

ALCOHOL DEPENDENCE AND CO-OCCURRING SUBSTANCE USE

Drug use disorders frequently co-occur with alcohol dependence. Based on both community and national samples, individuals with alcohol use disorders compared to those without alcohol use disorders are more likely to abuse or be dependent on illicit drugs. The Epidemiological Catchment Area (ECA) Study found that individuals diagnosed with alcohol abuse or dependence were 35 times more likely than those not diagnosed to use cocaine, 13 times more likely to use opioids, and 6 times more likely to use marijuana and related drugs. From NESARC data published 20 years later, a concurrent drug use disorder was identified in 13.05% of respondents with a current alcohol use disorder. Among those with co-occurring alcohol and drug use disorders, 73.9% were male. The most prevalent comorbid drug use disorders were cannabis (9.89%), cocaine (2.51%), and opioids (2.41%).

The NESARC study found that substance use and the related consequences of use are more severe when alcohol and other drugs are abused in combination than when abused alone. For example, the abuse of other drugs among alcohol-dependent persons is associated with poor treatment outcomes. Brown and his colleagues found at 6 months post substance abuse treatment completion that higher incidences of relapse and overall alcohol and drug use were reported by dual alcohol–cocaine users compared to alcohol-only users. Both groups showed improvement in functioning along legal, social, psychological, alcohol and drug dimensions at follow-up. Hoffmann and Miller found intravenous drug use and a history of antisocial behavior to be important predictors of relapse. Among 351 patients in Iceland, who sought alcohol detoxification, readmission during a 28-month follow-up period was associated with polysubstance abuse.

Patkar investigated the relationship between alcohol use and medical symptoms among cocaine-dependent patients, and found that patients who used alcohol had more respiratory, cardiovascular, digestive, head/neck, eye, and general health symptoms than cocaine-dependent patients who were nonalcohol users. Individuals who abuse both alcohol and drugs are also more likely to engage in behaviors that place them at risk to contract HIV, the human immunodeficiency virus. Stein and colleagues identified alcohol consumption as an independent risk factor for needle sharing among needle exchange program participants. Twenty-eight percent (28%) of needle exchangers met DSM-IV criteria for alcohol abuse during the previous 6 months. The odds of needle sharing for persons with at risk drinking was 2.3 times and with alcohol abuse was 2.5 times that of persons who drank no alcohol or low-risk levels of alcohol. Stein et al. also found that needle sharing and drinking were likely to occur together. Drinking was determined to occur on 40% of days when needles, cotton, or cookers were shared. In addition, needle exchange participants were found to be 1.5 times more likely to share needles, cottons, or cookers on alcohol use days.

Many individuals with alcohol use disorders smoke, putting them at high risk for such tobacco-related health problems such as cancer, lung disease, and heart disease. McAlister and Edmundson reported that heavy alcohol use was correlated with cigarette smoking and other forms of personal neglect. According to the 2001–2002 NESARC, 24.9% of the general population are current cigarette smokers. The rate of nicotine dependence was greater for individuals dependent on alcohol (45.4%) than for those with alcohol abuse (25.5%), and three times greater than those in the general population

(12.8%). Like comorbid alcohol and drug dependence, co-occurring alcohol and tobacco use is associated with health consequences in treatment populations. Hurt followed patients over 10 years after inpatient treatment for alcohol and other drugs to determine the impact of tobacco on alcohol-related deaths. Among deceased subjects, 50.9% died of tobacco-related illness and 34.1% died from an alcohol-related cause.

In summary, comorbid substance use disorder is not only prevalent among alcohol-dependent persons, but is also associated with many other physical, psychiatric, and social problems.

ALCOHOL DEPENDENCE, TREATMENT, AND RECOVERY

Alcohol dependence is a chronic disorder and affected persons can experience repeated cycles of treatment, abstinence, and relapse. Studies have shown however, not all recoveries from alcohol dependence occur with treatment. The data from the 2001–2002 NESARC found that nearly half of those diagnosed with an alcohol use disorder, prior to 1 year before the interview, were in various stages of recovery ranging from abstainers to asymptomatic risk drinkers. Only one quarter of those in recovery ever sought help for their alcohol problems. Additionally, a 3-year follow-up of those in recovery (in 2001–2002) showed that nearly half of asymptomatic risk drinkers had a recurrence of alcohol use disorder symptoms, as did 27% of low-risk drinkers and 7.3% of abstainers.

Investigators have examined the characteristics of those who seek treatment. Findings indicate that being male, younger age, having a previous treatment experience, and greater psychosocial problems are associated with treatment-seeking behavior. Contextual factors, like having a social network that discourages drinking and encourages help seeking, are also associated with treatment-seeking behaviors. Multiple treatment-seeking episodes are an indicator of severe alcohol dependence. Findings from the Combining Medications and Behavioral Interventions (COMBINE) study characterized subjects without previous treatment as having a later age of onset of alcohol problems and less psychological and physical distress.

In addition to comorbid psychiatric disorders, several other factors correlate with relapse among alcohol-dependent persons. Adamson et al. reviewed an extensive English-language literature on treatment outcomes from 1977 to 2005. This study identified gender as a significant predictor of treatment outcomes, with female gender being associated with better outcomes. Conversely, age, marital status, education, and ethnicity were generally poor predictors of treatment outcomes. Social functioning, measured by employment and higher socioeconomic status, predicted better outcomes. Substance use–related factors were consistent predictors of treatment outcomes. In particular, alcohol-related self-efficacy and motivation to change were consistently associated with better treatment outcomes, while baseline alcohol consumption, dependence severity, treatment history, duration of alcohol misuse, and craving/impaired control were moderately associated with poor treatment outcomes. The authors speculate that the associations between several of the alcohol-related variables were dependent on the measures used. There was vast variability in those studies reviewed related to the types and number of subjects, instruments used, and duration of follow-up.

This literature shows that recovery from alcohol dependence can occur without formal treatment. However, male gender, younger age, previous treatment experience, and greater psychosocial problems predict treatment seeking for alcohol problems. Alcohol dependence for some affected persons is a chronic, relapsing disorder.

TYPOLOGICAL CLASSIFICATIONS OF ALCOHOL-DEPENDENT PERSONS

While alcohol dependence is generally defined as a single clinical diagnostic entity, persons diagnosed with as alcohol dependent are heterogeneous. Drinking patterns, age of onset, co-occurring physical and psychiatric disorders, severity of symptoms, and the course of alcohol dependence vary considerably across affected individuals. Further, different demographic, family history, and other environmental factors contribute to this heterogeneity of persons with alcohol dependence. These differences affect treatment-seeking behavior and treatment outcomes by moderating or mediating clinical manifestations. The typological classification of alcohol dependence can be clinically useful by providing predictive value regarding the life course of alcohol dependence. Efforts to classify or group alcohol-dependent persons go back to the late nineteenth century, but earlier classifications were mostly based on drinking styles and the characteristics (e.g. chronicity and severity) of alcohol dependence. These approaches generally lacked predictive validity, while multidimensional approaches provided a more accurate description of the heterogeneity in alcohol-dependent persons.

Cloninger and colleagues first subtyped alcohol-dependent persons using multidimensional factors among 862 Swedish men adopted by nonrelatives at an early age. They found that both the genetic and the environmental backgrounds of the adoptees modified

their risk for alcohol abuse. Two forms of alcohol dependence were identified. Type 1 was characterized by mild or severe alcohol abuse and no paternal criminality. Type 2 was associated with severe and violence-related alcohol dependence. Broken down by heritability, Type 1 alcohol dependence tended to be influenced by environmental factors, while Type 2 was highly heritable from father to son. The genetic base of the Cloninger typologies has some support, but additional tests done in samples from Australia, the United States, and Europe resulted in mixed findings. Sannibale and Hall examined Cloninger's typology on 300 alcohol-dependent men and women in Australia. The Cloninger's typology failed to classify two types of problem drinkers and it did not predict gender differences in symptoms of alcohol dependence, family history or personality.

In the United States, Babor and colleagues conducted a cluster analysis of 321 male and female patients hospitalized for the treatment of alcohol dependence. The analysis was based on 17 defining characteristics including premorbid risk factors, alcohol and other drug use patterns, the chronicity and consequences of alcohol dependence, and comorbid psychiatric disorders. They derived two types of alcohol dependence, Type A and B. Type A was similar to the Cloninger's Type 1, and was characterized by a later onset of the disorder, fewer premorbid risk factors, and less psychopathology. Type B, similar to Type 2 was characterized by high-risk factors, early age of onset of alcohol dependence, familial alcoholism, and a more severe form of alcohol dependence. Babor et al.'s typology was replicated by outpatient samples of alcohol-dependent males and COGA samples. The relative proportion of Type A alcoholism is higher in the outpatient samples than in the inpatient samples. While Type A and Type B were similar to the Cloninger's classifications conducted on male only, the Babor's sample included both men and women. However, gender differences were not examined. The study of adolescents with alcohol use disorder found gender differences in two subtypes. Young women were more inclined toward affective disturbance, while young men tended toward conduct disorder.

Epstein and colleagues compared typological systems from five treatment studies, including (1) early versus late age of onset, (2) alcohol dependence with or without ASPD, (3) Cloninger et al.'s Type 1 and 2, and (4) Babor et al.'s Type A and B. They found that Type 1 and 2 classifications had poor construct validity due to symptom overlap. Stronger associations were found for early/late onset with Type 1 and 2 than with the Type A and B classifications; ASPD/Non-ASPD was associated with Type A/B classifications. Epstein and colleagues' study showed that while dichotomous subtypes can be easier to use in clinical situations and are conceptually more elegant, they may not be sufficiently complex to describe and adequately capture heterogeneous samples. Many older studies dropped cases that did not conveniently fit into two clusters or changed the cluster solution parameters in order to include the maximum amount of cases. Consequently, two-group solutions do not fully capture either clinical or general population samples. Depending on the variables of interest and the number of subjects examined, more recent studies have often identified three to five subtypes of alcohol dependence. The variability in the number of subtypes could be influenced by the data reduction technique used (cluster analysis, factor analysis, etc.), analysis not governed by proscribed rules, sample characteristics and sample sizes, the availability of clinical information, and the theory underlying the original analysis.

The indeterminate nature of cluster-derived typologies (and a limit of the statistical procedure) is best exemplified by a reanalysis of the Babor et al.'s data by Del Boca and Hesselbrock. The results showed four clusters as functional solutions that distinguished alcohol-dependent persons along gender and several clinically important dimensions. The largest subtype, containing approximately one-third of the cases was characterized as relatively Low Risk and Low Severity, including a mild form of alcohol dependence, late onset without family history, and low comorbid psychopathology. The High Risk/High Severity cluster included 22% of the female and 22% of the male patients, characterized by a severe form of alcohol dependence, early onset of alcohol use and dependence, positive family history of alcohol dependence, high alcohol involvement, conduct problems, polydrug use, depression, and ASPD. There were no gender differences in terms of the proportions or characteristics of subjects among both the mild and the severe forms of alcohol dependence. The two remaining clusters can be characterized as having moderate forms of alcohol dependence, labeled as Internalizing and Externalizing subgroups. There were gender-specific differences in both the Internalizing and the Externalizing subgroups. The Internalizing subtype included a higher proportion of women than men. This subgroup was characterized as depressed, anxious, and having severe alcohol dependence. They had medical or physical problems resulting from alcohol use, but a moderate family history risk. The fourth group, Externalizing subtype was predominantly male, was characterized as having a moderate family history risk, high levels of alcohol use, social consequences, and antisocial personality, but no depression or anxiety disorder.

Windle and Scheidt identified four cluster groups among inpatient subjects that have some similarity to those of Del Boca and Hesselbrock, including mild course, polydrug use, negative affect, and chronic/antisocial. Lesch and Walter also identified four alcohol

dependence types based on drinking patterns, severity, associated behavior, family history, and other risk factors in the sample of alcohol-dependent persons in Austria. Unlike US typologies, Lesch and Walter include motivations for drinking as defining characteristics. The Type 1 alcoholics, who drink and seek treatment for withdrawal, are characterized by a severe form of alcohol dependence and a positive family history without criminality. Type 2 alcoholics, drink to self-medicate for anxiety and conflict, and are characterized by low severity and no family history of alcoholism. Type 3 alcoholics who also self-medicate for depression are characterized by moderate severity, comorbidity, and a family history of psychiatric disorder, while Type 4 is characterized by premorbid brain injury, poor tolerance, and severe psychiatric, somatic and social problems.

Typological classifications have also been examined in large community samples. Johnson and his colleagues analyzed 911 Caucasian men in the ECA Survey and identified three subtypes along a severity dimension of alcohol dependence. The severe subtype was substantially associated with genetic influences, while the mild and dissociative subtypes were not. Furthermore, subjects in the severe subtype had significantly greater comorbid drug dependence and were more likely than other types to have sought treatment for alcohol problems. They suggest that these subtypes may be useful in identifying vulnerability genes in genetic linkage/association studies and help to identify environmental influences in behavioral and epidemiological studies. Johnson and colleagues reported similar findings among 3449 caucasian men and women with DSM-IV alcohol dependence who participated in the National Longitudinal Alcohol Epidemiologic Survey (NLAES).

More recently, Moss and colleagues identified five clusters among alcohol-dependent subjects in the NESARC study. Cluster 1 was characterized as a young adult subtype, associated with the beginning of their alcohol dependence career; Cluster 2 was a functional subtype, with late onset, low comorbidity, and low risk. Cluster 3 was an intermediate, familial subtype with late age of onset, high risk, moderate externalizing disorder, and high comorbid depression and anxiety disorder. Cluster 4 was labeled as a young antisocial subtype, characterized by an early age of onset, multigenerational family history, antisocial personality disorder, and the presence of depression, obsessive compulsive disorder (OCD), and substance abuse. Cluster 5 was labeled as a chronic severe subtype, characterized by late onset, moderate ASPD, the highest number of family history positives, and the highest level of psychiatric comorbidity.

While alcoholic typologies appear to capture complex symptom patterns, risk factors, and related conditions, it is difficult to synthesize these findings to be clinically useful. The variability of typology studies in the number and types of samples and the measurement and analytic approaches used make the application of their findings to clinical situations difficult.

A major limitation of multivariate typologies in the clinical setting is that most typologies contain too many defining characteristics, thus requiring a lengthy clinical assessment. Few treatment facilities are willing to devote additional time or personnel to obtaining the information necessary to permit reliable typological categorization of patients. Consequently, in order to increase their clinical utility, most typologies need to identify a limited number of critical indicators that can be readily identified, even if other features may have some theoretical importance. Regardless, the subtyping of alcohol-dependent persons may provide important information about the long-term course of alcohol dependence.

While many studies of alcoholic typologies do not have long-term follow-up or treatment outcome data, subjects in the Del Boca and Hesselbrock's study completed follow-up interviews. At 1-year and 3-year follow-up, the High Risk/High Severity and Externalizing groups reported poor alcohol-related outcomes. Twenty-five year mortality was also examined. A standard mortality ratio (SMR) was calculated for each cluster by gender. For men, the High Risk/High Severity cluster had the highest SMR despite being much younger than other cluster groups at admission, followed by the Externalizers. The Low Risk/Low Severity group and Internalizers had similar and lower SMRs. The number of women was too few to produce meaningful results.

The clinical application of alcoholic typologies based on two-group classifications (i.e. Type A/B and Type 1/2) in treatment studies have also been examined with mixed results. Litt and colleagues examined the usefulness of typological classification on predicting treatment outcomes. Seventy-nine empirically derived Type A and Type B alcohol-dependent males were randomly assigned to either coping skills training or interactional group psychotherapy. Interactional therapy was effective with Type A patients, while coping skills therapy was more effective with Type B patients. However, in a subsequent study with a much larger sample, the Type A/B classification did not produce a treatment matching effect with different forms of psychotherapy.

Typology-based pharmacological treatment studies of alcohol dependence have also produced mixed results. Dundon and colleagues found that Type A and B alcoholics had a differential response to pharmacotherapy with sertraline. Among Type A alcohol-dependent persons, those who received sertraline had better outcomes compared to those who did not while the

medications had no effect among Type B persons. Separate studies in Germany and Spain classified patients who received naltrexone treatment into Cloninger's Type 1 and 2. In both studies, naltrexone was no more effective than placebo among Type 1 patients with regard to the time to relapse. However, among Type 2 patients, naltrexone was effective in extending the time to relapse as compared to placebo. In contrast, a multisite study of fluvoxamine conducted in England, Australia, the United States, and Europe found that Type 2 patients were doing worse on fluvoxamine than on placebo 1 year after the treatment. No differences between fluvoxamine and placebo were observed among Type 1 patients. In a study of outpatients seeking treatment for alcohol dependence, the response to ondansetron was better predicted by the age of onset of alcohol dependence rather than Type A/B classifications.

The clinical literature thus far suggests an importance for typological classifications of alcohol dependence for clinical purposes, but it is not clear whether the Type 1/2 or A/B classifications are the most appropriate. The expanded 3–5 cluster typologies may be more clinically accurate for alcohol dependence, but their application in treatment studies is also more complex. Further, possible gender differences in response to pharmacotherapy should be considered in future studies.

CONCLUSIONS

This chapter reviewed alcohol dependence and several related conditions in adulthood. An early age of onset of alcohol use is associated with the early development of alcohol dependence (typically in early adulthood) and a severe course of alcohol dependence. While alcohol-related health conditions tend to be the result of chronic heavy drinking, alcohol-related psychiatric comorbidity could be both a risk factor for, and a consequence of, heavy drinking. Regardless of the nature of the association, psychiatric comorbidity consistently has a negative effect on the course of alcohol dependence.

Alcohol dependence in adulthood is a very heterogeneous disorder in terms of its etiology and clinical phenomenology. The nature of the associations among various alcohol-related conditions are also complex. There have been many attempts to identify more homogeneous subtypes of alcohol dependence. This chapter included a discussion of multivariate typological classifications of alcohol-dependent persons that take into account risk factors, alcohol use patterns, and related conditions. While no definitive typological classification of alcohol dependence is consistently found across clinical and community samples, these classification schemes appear to predict treatment responses, long-term outcomes, and mortality. More research efforts are needed to determine the number and types of classes that fully capture all cases of alcohol use disorder and are clinically useful.

SEE ALSO

Binge Drinking, The Biopsychosocial Model of Addiction, Disease Model, Epidemiology of Addiction, Gender Differences, Maturing Out, Natural Recovery, Relapse and Lapse, Self-Medication, Adolescent Substance Use: Symptoms and Course, Symptoms and Course: Older Age and Substance Abuse, Models of Relationships between Substance Use and Mental Disorders, Impact of Substance Use on the Course of Serious Mental Disorders, Substance Use and Mood Disorders, Substance Use in Response to Anxiety Disorders

Glossary

ASPD antisocial personality disorder.

Age of onset refers to the age that an individual first experiences a condition or symptom(s). In the study of addictions, the term is used for the age of first use of a substance, regular use, and the first symptom of dependence or abuse.

COGA Collaborative Study on the Genetics of Alcoholism. Funded since 1989 by the National Institute on Alcoholism and Alcohol Abuse (NIAAA) and more recently with contributions from the National Institute on Drug Abuse (NIDA), with the goal of identifying the specific genes underlying the vulnerability to alcohol dependence. The COGA investigators have collected extensive clinical, neuropsychological, electrophysiological, biochemical, and genetic data, and established a repository of immortalized cell lines from these individuals, to serve as a permanent source of DNA for genetic studies.

DSM the Diagnostic and Statistical Manual of Mental Disorders (DSM) is published by the American Psychiatric Association (APA) and provides a common language and standard criteria set for the classification of mental disorders. It is used in the United States and in varying degrees around the world by clinicians, researchers, psychiatric drug regulation agencies, health insurance companies, pharmaceutical companies, and policy makers. DSM has been revised several times since 1968, and currently DSM-IV has been used by clinicians and researchers. DSM-V is being field tested and is expected to be available in 2012.

The National Institute of Mental Health Epidemiological Catchment Area Survey (ECA) a comprehensive community-based survey of mental disorders and use of services by adults, ages 18 and older from 1980 to 1983. The survey was conducted in five communities in the United States through direct interviews or by proxy with 18 571 household and 2290 institutional residents using a standardized structured interview, the National Institute of mental Health Diagnostic Interview Schedule (NIMH-DIS).

NESARC the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) was designed to determine the magnitude of alcohol use disorders and their associated disabilities in the US general population. It is a longitudinal survey with its first wave of interviews fielded in 2001–2002 and a second wave in 2004–2005. The NESARC collects data on background characteristics, alcohol

and drug use, related conditions including psychiatric conditions, and family history information.

NSDUH the National Survey on Drug Use and Health (NSDUH) is conducted annually by the Substance Abuse and Mental Health Services Administration's Office of Applied Studies (SAMHSA; OAS). The survey provides information on the prevalence, patterns, and consequences of alcohol, tobacco, and illegal drug use and abuse in the general US civilian noninstitutionalized population, age 12 and older.

Standard Mortality Ratio (SMR) the ratio of observed deaths to expected deaths according to a specific health outcome in a population and serves as an indirect means of adjusting a rate.

Telescoping effect refers to the accelerated progression of the course of substance use and dependence (i.e. from the age of first use/regular use to dependence, treatment entry, etc.)

World Health Organization (WHO) the directing and coordinating authority for health within the United Nations system. It is responsible for providing leadership on global health matters, shaping the health research agenda, setting norms and standards, articulating evidence-based policy options, providing technical support to countries, and monitoring and assessing health trends.

Further Reading

Adamson, S.J., Sellman, J.D., Frampton, C.M., 2009. Patient predictors of alcohol treatment outcome: a systematic review. Journal of Substance Abuse Treatment 36, 75–86.

Babor, T.F., et al., 1992. Types of alcoholics, I. Evidence for an empirically derived typology based on indicators of vulnerability and severity. Archives of General Psychiatry 49, 599–608.

Chick, J., Aschauer, H., Hornik, K., 2004. Efficacy of fluvoxamine in preventing relapse in alcohol dependence: a one-year, double-blind, placebo-controlled multicentre study with analysis by typology. Drug and Alcohol Dependence 74, 61–70.

Cloninger, C.R., Bohman, M., Sigvardsson, S., 1981. Inheritance of alcohol abuse. Cross-fostering analysis of adopted men. Archives of General Psychiatry 38, 861–868.

Corrao, G., et al., 2004. A meta-analysis of alcohol consumption and the risk of 15 diseases. Preventive Medicine 38, 613–619.

Dawson, D.A., Goldstein, R.B., Grant, B.F., 2007. Rates and correlates of relapse among individuals in remission from DSM-IV alcohol dependence: a 3-year follow-up. Alcoholism: Clinical and Experimental Research 31, 2036–2045.

Del Boca, F.K., Hesselbrock, M.N., 1996. Gender and alcoholic subtypes. Alcohol Health and Research World 20, 56–62.

Epstein, E.E., et al., 2002. A multi-site study of alcohol subtypes: classification and overlap of unidimensional and multi-dimensional typologies. Addiction 97, 1041–1053.

Grant, B.F., et al., 2004. Prevalence, correlates, and disability of personality disorders in the United States: results from the national epidemiologic survey on alcohol and related conditions. Journal of Clinical Psychiatry 65, 948–958.

Hesselbrock, V.M., Hesselbrock, M.N., 2006. Are there empirically supported and clinically useful subtypes of alcohol dependence? Addiction 101 (Suppl. 1), 97–103.

Johnson, E.O., et al., 1998. Extension of a typology of alcohol dependence based on relative genetic and environmental loading. Alcoholism: Clinical and Experimental Research 22, 1421–1429.

Lesch, O.M., Walter, H., 1996. Subtypes of alcoholism and their role in therapy. Alcohol and Alcoholism 31 (Suppl. 1), 63–67.

Moss, H.B., Chen, C.M., Yi, H.Y., 2007. Subtypes of alcohol dependence in a nationally representative sample. Drug and Alcohol Dependence 91, 149–158.

Mueser, K.T., et al., 2006. Comorbid substance use disorders and psychiatric disorders. In: Miller, W.R., Carroll, K.M. (Eds.), Rethinking Substance Abuse: What the Science Shows, and What We Should Do About It. Guilford Press, New York, pp. 115–133.

Rehm, J., et al., 2003. Alcohol-related morbidity and mortality. Alcohol Research and Health 27, 39–51.

Research Society on Alcoholism 2011. Impact of alcoholism and alcohol induced disease on America http://www.rsoa.org/2011-04-11RSAWhitePaper.pdf.

Stein, M.D., et al., 2002. Alcohol and HIV risk taking among intravenous drug users. Addictive Behaviors 27, 727–736.

Stinson, F.S., et al., 2006. Comorbidity between DSM-IV alcohol and specific drug use disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Alcohol Research and Health 29, 94–106.

Windle, M., Scheidt, D.M., 2004. Alcoholic subtypes: are two sufficient? Addiction 99, 1508–1519.

Relevant Websites

http://www.aaap.org/ – American Academy of Addiction Psychiatry.
http://www.dsm5.org/ – American Psychiatric Association DSM-V.
http://www.apa.org – American Psychological Association.
http://www.asam.org – American Society of Addiction Medicine.
http://samhsa.gov/about/csap.aspx – Center for Substance Abuse Prevention.
http://samhsa.gov/about/csat/aspx – Center for Substance Abuse Treatment.
http://www.naadac.org – NAADAC Substance Abuse Professionals.
http://www.helpstartshere.org – National Association of Social Workers.
http://store.samhsa.gov/home – National Clearinghouse for Alcohol and Drug Information.
http://www.niaaa.nih.gov – National Institute on Alcohol Abuse and Alcoholism.
http://www.nida.nih.gov – National Institute on Drug Abuse.
http://www.nimh.nih.gov – National Institute of Mental Health.
http://www.RSOA.org – Research Society on Alcoholism.
http://findtreatment.samhsa.gov – Substance Abuse Treatment Facility Locator.

CHAPTER

12

Symptoms and Course: Older Age and Substance Abuse

Catherine A. Haighton[*], *Graeme B. Wilson*[*], *Jonathan Ling*[$], *Karen E. Smith*[$], *Ann Crosland*[$], *Eileen F.S. Kaner*[*]

[*]Newcastle University, Newcastle upon Tyne, UK, [$]Sunderland University, Sunderland, UK

OUTLINE

INTRODUCTION

Positive healthy behaviors can not only extend longevity but also reduce the risk of losing mobility and independence in later life, and there is already a body of evidence to show that lifestyle behavior differs according to age, with older people making up a large but unique target population for interventions. However, substance abuse – the use and misuse of illicit and/or licit substances – does occur in older age. Problems arise from drinking excessive levels of alcohol, smoking tobacco, misuse of prescribed medications or over-the-counter medication, and illicit drug use. While some substance use decreases with age (e.g. alcohol) other substance use is more common in later life due to social and generational effects (i.e. smoking was more acceptable when many older people began). Also, with increasing health problems in later life, prescription drug use is more common and this can become problematic due to confusion and/or dependence-forming properties. Lastly, use of illicit drugs is relatively uncommon in later life, but their addiction and criminogenic effects can present a significant problem for older individuals. Substance abuse is associated with notable social, psychological, physical, and economic consequences. While older age substance abuse is often less obvious, misdiagnosed, undetected or under-detected, or under-reported, a recent survey predicts that the number of Americans aged 50 or older who are addicted to drugs or alcohol will double by the

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00012-7

year 2020. The aging of populations worldwide means that the absolute number of older people with substance abuse problems is on the increase and a real danger exists that a "silent epidemic" may be evolving.

The question of what constitutes old age is a vexed one. The usual cutoff point for old age in the West is 65 years of age. The United Nations defines the older population as those over 60 years, but recently it has modified its definition for older people in Africa to include those over 50 years in order to take into account environmental conditions. There is little consensus within the academic literature regarding the definition of an older person and although this term is widely used, definitions vary from 50 years upwards. In order to be as inclusive as possible with the literature on older age and substance abuse individuals aged 50 years or more are included.

ALCOHOL

Extent of Use

General population figures indicate that alcohol use decreases with age. However, since 2001, surveys have indicated a small but steady increase in the amount of alcohol consumed by the middle and older age groups. Men and women aged 65 and over also drink more frequently than those in younger age groups.

Drinking habits of middle-aged and older adults in America are broadly similar to the use of alcohol in Canada and Brazil with 19% of males aged 50–64, and 13% of females drinking at risky levels. In those aged 65 or over, 13% of males and 8% of females report at-risk alcohol use (more than two drinks per day). In the 50–64 age group, 23% of males and 9% of females also report binge drinking (defined as five or more drinks on the same occasion), and 14% of men and 3% of women aged 65 years and over report binge drinking. Older women may be especially at-risk for alcohol problems because they are more likely to outlive their spouses and face other losses that may lead to loneliness and depression. Physiologically, women are also at greater risk as they have a greater predisposition to have less lean muscle mass than men, making them more susceptible to the effects of alcohol. Both men and women experience losses in lean muscle mass as they age, but women have less lean muscle mass than men throughout adulthood and, therefore, are less able to metabolize alcohol throughout their lives, including into older adulthood. In addition to gender, having a college degree, being separated, having a middle or high income, smoking and drug use are all associated with at-risk and binge drinking. These findings have been replicated in other studies of older people from both developed (USA, France, Korea, Germany) and developing (Brazil) countries.

In the United Kingdom, 16% of men and 5% of women aged 65 and over drink more than the daily recommended limits (4 units day^{-1} for men and 3 units day^{-1} for women) while 15% of men and 7% of women aged 65 and over and 23% of men and 8% of women aged 60 and over drink more than the recommended weekly limits (21 units $week^{-1}$ for men and 14 units $week^{-1}$ for women). Also, in the United Kingdom, 31% of 50–59 year olds and 22% of 60–69 year olds report hazardous drinking behavior in the preceding 12 months, as measured by the Alcohol Use Disorders Identification Test (AUDIT). The 2007 Eurobarometer survey estimates that 27% of European people aged 55+ years had episodes of binge drinking (5+ drinks of 50 g alcohol on a single occasion) at least once a week during the previous 12 months. Those aged 50–65 years have the highest weekly expenditure on alcohol. There is also a substantial prevalence of alcohol problems in older hospital patients and in nursing homes with the possibility that more than one quarter of nursing home residents may have symptoms of active problem drinking.

Effects of Use

Negative

Excessive alcohol consumption is strongly associated with serious social, psychological, physical, and economic costs for the individual and society. As people age, physiological changes that occur as part of the aging process mean that older people are more sensitive to the effects of alcohol and experience problems at lower levels of consumption. The average person aged 65 years and over also takes at least two medicines per day some of which can be dangerous when mixed with alcohol. Combined use of alcohol and medication is estimated to affect up to 19% of older Americans although there is some evidence that older people may cease or reduce drinking as a result of changes in medication use or physical illness. Drinking alcohol for medicinal purposes is also common and some people, especially older women, may find it easier to discuss their alcohol consumption in the context of medicinal use. Older adults with more health problems and who rely on alcohol to manage pain are at elevated risk for drinking problems. The presentation to health professionals of older people with alcohol problems may be atypical (such as falls, acute hip fracture, neglect, malnutrition, depression, or confusion which can be mistaken for Alzheimer's) or masked by comorbid physical or psychiatric illness which makes detection all the more difficult. Some medical conditions (such as high blood pressure, ulcers, and diabetes) can worsen with alcohol

use. For people aged over 60 years, drinking more than seven drinks per week or more than three drinks per occasion has been associated with impairments in the instrumental activities of daily living. Men aged 55–64 years and women aged 65–74 years have the highest death rates from liver disease, while the highest rates of deaths each year, directly attributable to alcohol, occur in the 55–74 age group.

Positive

In moderation, alcohol consumption can contribute to older peoples' quality of life. Benefits of low-to-moderate alcohol consumption for older people include a lower risk of Alzheimer's disease, dementia, and functional decline, a reduced risk of stroke, protection against heart disease and respiratory disease and lower overall mortality. Older men and women who consume small to moderate amounts of alcohol are more likely to maintain mobility than nondrinkers. Light to moderate alcohol use by older women has also been associated with better self-perceived health status and lower rates of hospitalization. In middle-aged and older men and women, moderate levels of alcohol consumption are associated with better cognitive health than abstinence and moderate consumption of alcohol may increase bone mineral density in men and postmenopausal women. However, a consumption increase is unlikely to produce any net benefit in terms of cardiovascular mortality when lifestyle factors are taken into consideration – at least in older drinkers. Recent estimates are such that the potential health benefits of alcohol are significantly outweighed by the harm caused.

Course

Early-onset drinkers or "survivors" are those with a continuing problem with alcohol which developed in earlier life. It is thought that two-thirds of older problem drinkers have had an early-onset of alcohol misuse. These drinkers are likely to need more intensive support than brief interventions. Late-onset drinkers or "reactors" begin problematic drinking later in life, often in response to traumatic life events such as bereavement, loneliness, pain, insomnia, or retirement. This population is more likely to be female. As "reactors" also have higher life satisfaction, less physical and psychological comorbidity, and fewer alcohol-related problems than "survivors," they are often difficult to identify and diagnose. Evidence also suggests that as a population late-onset drinkers are more motivated to change their drinking behavior, more compliant in treatment and have better treatment outcomes. Intermittent or binge drinkers are defined in terms of consumption thresholds of six standard drinks for females or eight for males, in a single day. Identifying heavy episodic consumption in this population is made difficult in part because the usual cues (employment-related difficulties and driving offenses) are more rare in the older population – many of whom do not work or drive. Like late-onset drinkers, this group will be likely to respond well to alcohol treatment or interventions.

Although there is growing recognition of this public health problem and instruments to detect problem drinking generally perform well in older populations, problems have been identified when screening and intervening for alcohol misuse including inaccurate self-report, staff embarrassment, poor use of the screening tool by professionals, and inaccurate patient history. Some of these factors may be more influential when screening an older population given a widespread attitude that alcohol is a problem particular to younger age groups. Greater social stigma attached to older people's drinking may exacerbate under-reporting or nonreporting of alcohol problems or levels of drinking particularly among older women, who have been found to be less likely to present for treatment for alcohol problems than older men. Professionals can be reluctant to ask older people about alcohol use due to awkwardness or stereotyped views about older people and people who have alcohol problems; and alcohol problems are often presented in nonspecific ways among older people. Even when clinicians have been encouraged to screen for problem alcohol use, they under-deliver health-promoting advice to older people, while nurses report that they do not engage with older people as they worry about depriving them of the social benefits of drinking. Only a small proportion of older people with alcohol-related problems are referred on for rehabilitation as health care workers are less likely to refer older people for specialist treatment than they are to refer younger people.

Limited evidence exists for the most efficacious approaches to prevent and/or intervene for alcohol problems in older people. Many strategies are neither age-specific nor sensitive to what is most clinically effective in accommodating the unique biological and social conditions of older persons. However, older people can benefit as much from treatment for alcohol misuse as younger age groups, and older people with dependence have been found to respond better to treatment than younger age groups. Those whose alcohol problems are of late-onset have fewer health problems and are more receptive to treatment than those with early-onset problems associated with previous family history; they are also more likely to recover spontaneously. Older adults have better attendance at, and adherence to, therapy and medications than younger age groups.

Older people are, however, one of the least well-informed groups about alcohol units and there is little evidence regarding effectiveness of education campaigns.

Brief interventions and cognitive behavioral approaches have been shown to be effective at reducing alcohol consumption among older adults. Simple brief interventions, typically involving a structured advice session of a few minutes with a trained practitioner (not necessarily a specialist), are effective in reducing hazardous drinking. Extended brief interventions, involving 20–30 min of structured therapy possibly with follow-up sessions, are efficacious with those for whom a short brief intervention is not successful, or who have experienced harm in relation to drinking alcohol. Brief interventions are seen as an acceptable and effective treatment in reducing hazardous drinking among older adults. Inpatient (rather than outpatient) detoxification is recommended for older people who are severely alcohol dependent although information on the use of abstinence medication in older people is limited. Interventions that involve or treat family members or encompass an individual's support network have been endorsed, particularly as a means of countering relapse for older people with alcohol problems.

Rather than a single treatment of choice, a range of treatment options may be beneficial at different levels of drinking: low-risk, hazardous, harmful, moderate dependency, and severe dependency. Groups with complex needs are likely to require more intensive or prolonged interventions, even at lower levels of alcohol use and dependence. A stepped care approach is recommended as a practical and efficient alternative to one-size-fits-all recommendations.

Some experts suggest that recommended limits for intake should be redefined for older people. In the USA, the National Institute on Alcohol Abuse and Alcoholism recommends that people over 65 years of age should consume no more than one standard drink per day, seven standard drinks per week, and no more than two drinks at any one time. Among the member states of the European Union only Italy has an alcohol consumption guideline for older adults. The recommendation is no more than one standard drink per day.

TOBACCO

Extent of Use

One out of five adults over the age of 50 years smokes cigarettes. Today's generation of older Americans had smoking rates among the highest of any USA generation. In the mid-1960s, about 54% of adult males were current smokers and another 21% were former smokers; in 2008, about 23% of adult males were smokers and another 24% were former smokers. In 2008, over 17 million Americans over the age of 45 smoked, accounting for over 22% of all adult smokers. Nine percent of Americans over 65 years of age currently smoke. Older people have usually smoked for longer than younger people and have been and continue to be heavy smokers. Older smokers also smoke more than younger people and are more likely to smoke brands of cigarettes that have high nicotine levels. In the United Kingdom, the highest rates of smoking are in the 20–24 age group, with 30% of people this age being recorded as smokers. The prevalence of smoking then declines with age to 13% of people aged 60 and over smoking. The difference between the age groups has historically been smaller and has increased as a result of higher smoking cessation rates among older people. For example, in 1974, 34% of people aged 60 and over smoked. By 2008, this had more than halved to 13%, whereas the decrease for people aged 20–24 over the same time period is from 48 to 30%. In the United Kingdom, the over 65s, supported by National Health Service (NHS) smoking cessation services, are giving up smoking more quickly than any other age group. Unlike other drugs, tobacco is used openly and therefore there should be no difficulty with recognizing its use. However, in an Australian study of hospital patients over the age of 65 years, 12% were smoking a median of 17 cigarettes per day (range 1–50), only 38% of whom had been identified as smokers by ward doctors.

Effects of Use

Tobacco is the most commonly used psychoactive substance in older people and accounts for more medical disability and mortality than any other substance of potential misuse. Older smokers are more likely to have chronic diseases than nonsmokers with their smoking causing further deterioration in their health. Tobacco use has been shown to cause about 25 life-threatening diseases and smoking is a major risk factor in eight of the top sixteen causes of death in people aged 65 years and older. It is well documented that smoking is directly responsible for more than 90% of chronic obstructive pulmonary disease (COPD, or emphysema and chronic bronchitis) deaths and approximately 80–90% of lung cancer deaths in women and men, respectively. Smoking is also a major risk factor for coronary heart disease, stroke, and lower respiratory tract infections – all leading causes of death in those over 50 years of age. Older smokers are at greater risk from smoking because they have smoked longer (an average of 40 years), tend to be heavier smokers, and are more likely to suffer from smoking-related illnesses. They are also significantly less likely than younger smokers to believe that smoking harms their health. COPD prevalence rates are highest among those 65 years of age and older and the disease consistently ranks among the top

10 most common chronic health conditions and sources of daily activity limitation. COPD is the fourth leading cause of death and is predicted to become the third by 2020. Men aged 65 years and over who smoke are twice as likely to die from a stroke, and women smokers are about one and a half times as likely to die from a stroke, as their nonsmoking counterparts. The risk of dying from a heart attack is 60% higher for smokers than nonsmokers among those aged 65 years. Smokers are more than twice as likely as nonsmokers to develop dementia of any kind and Alzheimer's disease. Smokers also have two to three times the risk of developing cataracts, the leading cause of blindness and visual loss, as nonsmokers. Smoking reduces a person's normal life expectancy by an average of 13 to 15 years, thereby eliminating or shortening retirement years for most smokers. The complications of smoking occur later in the period of use and it may therefore appear that a person has smoked for many years without apparent ill effect.

Course

Although quitting smoking has proven health benefits, even at a late age, older smokers (50–74) are less likely to have tried to quit than smokers aged 21–49. When an older person quits smoking, circulation improves rapidly, and the lungs begin to repair damage. In one year, the added risk of heart disease is cut almost in half, and risk of stroke, lung disease, and cancer diminish. Many older adults say they do not quit smoking because doing so offers no benefit at an advanced age. However, there is strong evidence that smoking cessation even late in life not only adds years to life, but also improves quality of life. Treatments for quitting smoking have been found to be effective and could decrease health care costs. Using a tobacco treatment plan doubles the quitting success rate. Effective treatments combine counseling and medications. Older people have been shown to have greater success in stopping smoking than younger people with 60% of those aged 60 and over successfully quitting in 2007/08, compared with just 38% of those under 18. Smoking cessation is beneficial at all ages and has immediate and long-term benefits for those with and without smoking-related illnesses. About 75% of years of potential life lost in smokers occur after the age of 65. Former smokers live longer than continuing smokers as smoking cessation at age 50 reduces the risk of dying within the next 15 years by 50%.

Smoking remains prevalent in older people in the United Kingdom and similar countries. The adverse health effects of current smoking continue to accumulate in old age and stopping smoking in old age confers benefits on function, morbidity, and mortality. Many older people wish to stop smoking and many are successful, although cognitive impairment – more prevalent in this population – is a barrier to successful smoking cessation.

Smoking cessation interventions have been shown to be effective. Although self-help strategies alone marginally affect quit rates, individual and combined pharmacotherapies and counseling either alone or in combination can significantly increase cessation. Pharmacotherapies include all forms of nicotine replacement therapy (NRT) which can help people quit smoking, almost doubling long-term success rates. All forms of NRT make it more likely that a person's attempt to quit smoking will succeed. There is no evidence that one form of NRT is better than any other and NRT works with or without additional counseling. Using effective smoking treatments is strongly encouraged for all populations, especially those with high and heavy rates of smoking, such as psychiatric and substance abuse populations. Smokers of all ages benefit from cessation of smoking. Although most older smokers, like younger smokers, prefer to quit on their own, there are few interventions targeted at older smokers. A number of studies indicate that older smokers face age-specific obstacles to quitting, such as decreased self-efficacy due to greater number of lifetime quit attempts; longer smoking history and a tendency to be heavier, addicted smokers; and skepticism about the benefits of quitting. Moreover, a common misconception among lay and professional populations is that it is "too late" for older smokers. These data illustrate the specific smoking behavior of older people, suggesting that cessation interventions ought to be tailored to these characteristics. Indeed tailored interventions tend to be rated more highly and are also read and reread more than nontailored interventions. Quit rates are also higher among smokers who received a tailored intervention. Doctors and other health care staff should make use of patient contact to encourage older smokers to quit.

PRESCRIPTION/OVER-THE-COUNTER MEDICATION

Extent of Use

Older people are frequent and regular consumers of prescription medication (including diuretics, benzodiazepines, opioid analgesics, hypnotics, and antidepressants) and over-the-counter medication (such as laxatives, analgesics, and antihistamines). Persons aged over 65 years use about one-third of all prescribed medication. Older people receive more prescribed medications than younger people and are far more likely to be prescribed benzodiazepines on long-term prescriptions than younger people. This "medical" use of

dependence-creating pharmaceuticals is often related directly to the aging process and increased physical morbidity. Older women are more likely to be prescribed, and to misuse, psychoactive medications than men and are also at a higher risk of prescription medication misuse than other age groups. In the USA, up to 11% of older women misuse prescription drugs and prescription drug abuse is present in 12–15% of older individuals who seek medical attention. Inappropriate prescribing is common among community-dwelling older people and persists over time. Surveys of older community-based individuals indicate from 17 to 20% are receiving psychotropic medications that are highly addictive. Almost half indicate use beyond the prescribed levels of these medications. One quarter of the prescription medications sold in the USA are used by older people – often for problems such as chronic pain, insomnia, and anxiety. Prescription/over-the-counter medication abuse is associated with female gender, social isolation, recent divorce, widowhood, lower educational level, lower income, poorer health status, depression, anxiety, and history of substance abuse. An Australian study of hospital patients found 14% were problem users of benzodiazepines and had taken these drugs for periods ranging from 6 months to 20 years. Only 3% had been identified by ward doctors. Nonmedical use of drugs among adults older than 50 years is set to increase to 2.7 million by the year 2020.

Prescription of medicines is a fundamental component of the care of elderly people, and optimization of drug prescribing for this group of patients has become an important public health issue worldwide. Several characteristics of aging and geriatric medicine affect medication prescribing for elderly people and render the selection of appropriate pharmacotherapy a challenging and complex process. Some syndromes related to age, especially cognitive impairment, affect the ability of elderly people to engage with health services. For example, elderly people with dementia have increased difficulty with taking drugs, and dementia impedes their ability to make autonomous decisions about their medicines. In addition, frail elderly people have age-related impairments in the hepatic metabolism and renal clearance of medications, and enhanced pharmacodynamic sensitivity to specific drugs.

Elderly patients are also at high risk of having drug interactions, but the prevalence of these interactions is not well documented. Several types of interactions exist: drug–drug, drug–disease, drug–food, drug–alcohol, drug–herbal products, and drug–nutritional status. Factors such as age-related changes in pharmacokinetics and pharmacodynamics, frailty, individual variability, reduced homoeostatic mechanisms, and psychosocial issues need to be considered when drug interactions are assessed. Software can help clinicians to detect drug interactions, but many programs have not been updated with the evolving knowledge of these interactions, and do not take into consideration important factors needed to optimize drug treatment in elderly patients. Any generated recommendations have to be tempered by a holistic, geriatric, multiprofessional approach that is team based.

Effects of Use

Problems associated with prescription/over-the-counter medication abuse in older people are particularly likely to go unrecognized. The older abusers have a high rate of psychiatric problems due to addiction. Unlike younger addicts older misusers are more likely to have physical and neurological complications due to the addiction. Older adults are likely to experience more problems with relatively small amounts of substances because of increased sensitivity, slower metabolism, and a smaller volume of distribution. Age-related cognitive impairment may interfere with the ability of the older adult to self-monitor intake or interpret feedback from health care providers.

Course

Screening instruments for prescription drug abuse have not been validated in the older population. Benzodiazepines, opiate analgesics, and some skeletal muscle relaxants may result in physical dependence; however, tolerance, withdrawal syndrome, and dose escalation may be less common in the older patient. Lower doses may decrease the risk of abuse and dependence; however, fear of abuse often results in a failure to treat adequately symptoms such as anxiety, pain, and insomnia.

The treatment of disorders of prescription drug use in older adults may involve family and caretakers and should take into account the unique physical, emotional, and cognitive factors of aging. The use of nonaddicting pain treatment modalities, such as physical therapy, occupational therapy, acupuncture, nonaddicting medication, and support groups focusing on management of pain and disability are indicated in the context of cognitive behavioral therapy.

ILLICIT DRUGS

Extent of Use

The use of illicit substances does occur among older people but the research evidence concerning prevalence is equivocal; the Epidemiological Catchment Area

Survey (ECAS) reported a rate of illicit drug use of less than 1% in people over the age of 60 (1988), while in 1994 a study of US veterans reported a rate of 22% in those aged 55 and over. In the United Kingdom there has been a significant increase in the proportion of drug users aged 50 and over in contact with specialist drug treatment services (from 1.7% in 1998 to 3.5% in 2005). There is also a similar trend among those in contact with agencies that provide clean injecting equipment to drug users (syringe exchange schemes) with the median age of injectors in contact with such services increasing by almost 8 years over a 13-year period from 27.0 in 1992 to 34.9 in 2004. The UK's drug treatment services and syringe exchange schemes typically cater for drug-dependent people who are usually users of opiates (mainly heroin) or stimulants (cocaine, crack cocaine, and amphetamine) or who inject drugs. The European Monitoring Centre for Drugs and Drug Addiction has highlighted aging populations of opiate users in a number of European countries. In Great Britain, changes in the age of people taking illicit substances are monitored through the British Crime Survey (BCS). Against a backdrop of significant falls in the rate of last-year prevalence for the use of any illicit drug (predominantly cannabis, cocaine, ecstasy, amyl nitrate, and amphetamines) for the youngest age groups (16–29 years) between 1996 and 2008/09, the BCS shows that illicit drug use among those aged 30–59 years has increased. No data on illicit drug use among people aged over 60 is available from the BCS because it does not collect this information due to the perceived "very low prevalence rates for use of prohibited drugs" among people aged 60 and over, reflecting the prevalent attitude that older people do not use drugs. Like their younger counterparts, older adults do use illicit drugs and although illicit drug use is less common in this age group, its prevalence is increasing. Estimates from the USA suggest that the number of persons aged over 50 needing treatment for illicit drug problems may increase by up to 300% between 2001 and 2020. As methadone and other maintenance programs become more effective at retaining patients in treatment and reducing overdose deaths, the number of older patients will gradually increase although this is subject to the ongoing availability of such programs.

Effects of Use

Illicit drug use by older people presents unique problems as biological systems and processes alter naturally across the life course. However, the effect of concurrent drug use on some of these systems is not well understood. The natural progression of certain diseases means that symptoms manifest only in older age and the lives of older drug users are likely to be characterized by considerable levels of morbidity. The brain changes in a variety of ways across the lifespan, and how these changes alter drug–brain interactions and what implications these changes have for older drug users are not yet clear. In addition, chronic use of some drugs may exacerbate changes normally associated with aging; people dependent upon cocaine, for example, exhibit an increased number of age-related white matter (brain) lesions, which in turn are thought to be associated with cognitive abnormalities. Pharmacokinetics – the processes by which a substance is absorbed, distributed, metabolized, and eliminated from the body – also changes with age. Reductions in lean body mass and total body water content, coupled with reduced drug elimination by the kidneys, may increase elevated drug serum levels, and even moderate use of drugs may have significant effects. Long-term drug use further increases the risk of certain morbidities already prevalent in older age such as myocardial, pulmonary, and cerebral infarctions, which are associated with cocaine use. The natural progression of other diseases, for example, cirrhosis and other liver diseases (associated with hepatitis C infection contracted through the sharing of contaminated drug injecting equipment and/or excessive alcohol use), means that symptoms tend to manifest only in drug users of older age. Concurrent aging and drug use therefore create a unique set of problems for older people that are not yet fully understood.

Course

Historically, global populations have not witnessed a large number of older illicit drug users and this has resulted in a perception that older people do not use these substances. However, cross-sectional studies fail to account for period and cohort effects and the likelihood that older people in the past did not use drugs because they did not use them when they were younger. Older people of today are using drugs because they did so when younger, and have done little to change their consumption as they have aged. Whether there exists a second group of older users of illicit drugs – those who were abstemious when young but who commenced use in later life – remains unknown, because the lack of awareness of drug use among older populations has largely precluded any investigation of this issue and this has been compounded by the fact that the tools used to screen for drug use have not been validated for use in older populations. The advent of effective treatment and harm minimization initiatives for drug-dependent individuals in the past 30 years or so, in addition to general advances in medicine, has increased the average life expectancy of a drug user, and the trends described here demonstrate their survival into older

age. Despite positive evidence of the effectiveness of treatments for other substances, there is little substantive literature regarding the effectiveness of treatments for older illicit substance users.

CONCLUSION

Substance abuse does occur in older age and is associated with notable social, psychological, physical, and economic consequences. Older age substance abuse is often less obvious, misdiagnosed, undetected or under-detected, or under-reported. Screening tools aimed at identifying people at-risk of substance abuse are not designed with an older population in mind although older people seem to respond well to treatment. Older people are likely to perceive great social pressure not to be seen as a substance abuser, and this is likely to have an impact on self-reported statistics. The aging of populations worldwide means that the absolute number of older people with substance abuse problems is on the increase and a real danger exists that a "silent epidemic" may be evolving. The likely silence of many older people means that the epidemic could be greater than even this account suggests.

SEE ALSO

Alcohol Use Disorders, Heroin Addiction, Cocaine Addiction, Prescription and Over-the-Counter Medications, Tobacco, Binge Drinking, Gender Differences, Self-Medication

List of Abbreviations

AUDIT	Alcohol Use Disorders Identification Test
COPD	Chronic Obstructive Pulmonary Disease
NHS	National Health Service

Glossary

Abstinence a voluntary restraint from indulging in bodily activities that are widely experienced as giving pleasure.

Alcohol units a simple way of remembering alcohol strengths. There is one unit of alcohol in one half-pint of ordinary strength beer, larger or cider (284 ml, 3.5% ABV); one pub measure of spirits (25 ml, 40% ABV); one standard glass of wine (125 ml, 8% ABV); one small glass of sherry (50 ml, 20% ABV). Also referred to as standard drink.

Binge drinking defined as episodic excessive drinking. There is currently no worldwide consensus on how many drinks constitute a "binge," but in the United States, the term is often taken to mean consuming five or more standard drinks (male), or four or more drinks (female), in about 2 h for a typical adult. In the United Kingdom, binge drinking is defined as drinking more than twice the daily limit, that is, drinking eight units or more for men or six units or more for women.

Comorbidity the presence of one or more disorders (or diseases) in addition to a primary disease or disorder, or the effect of such additional disorders or diseases.

COPD Chronic Obstructive Pulmonary Disease

Criminogenic producing or tending to produce crime or criminality.

Dependent drinking when a person feels that they are unable to function without alcohol.

Detoxification the physiological or medicinal removal of toxic substances from a living organism, including, but not limited to, the human body and additionally can refer to the period of withdrawal during which an organism returns to homeostasis after long-term use of an addictive substance.

Harmful drinking when a person drinks over the recommended weekly amount and has experienced health problems directly related to alcohol.

Hazardous drinking when a person drinks over the recommended weekly limit or binge drinks. Individuals may not yet have any health problems directly related to alcohol but are increasing the risk of experiencing problems in the future.

Illicit not sanctioned by custom or law; unlawful.

Licit permitted by law; legal.

Mortality susceptible to death.

Pharmacokinetics the processes by which a substance is absorbed, distributed, metabolized, and eliminated from the body.

Psychoactive affecting the mind or mental processes.

Reactors or late-onset drinkers begin problematic drinking later in life, often in response to traumatic life events such as bereavement, loneliness, pain, insomnia, or retirement.

Relapse when a person is affected again by a condition that affected them in the past.

Substance abuse the use and misuse of illicit and/or licit substances.

Survivors or early-onset drinkers a continuing problem with alcohol which developed in earlier life.

Further Reading

Andréasson, S., Hallgren, M., Högberg, P., 2009. Alcohol consumption among elderly European Union citizens: health effects, consumption trends and related issues. In: Health SNIoP (Ed.), Expert Conference on Alcohol and Health. Sweden, Stockholm.

Barry, K.L., Blow, F.C., 2003. Use and Misuse of Alcohol among Older Women. National Institute on Alcohol Abuse and Alcoholism. http://pubs.niaaa.nih.gov/publications/arh26-4/308-315.htm.

Beynon, C., 2009. Drug use and ageing: older people do take drugs!. Age and Ageing 38, 8–10.

BMA Board of Science, 2008. Alcohol Misuse: Tackling the UK Epidemic. BMA, London.

Office for National Statistics, 2008. Drinking: Adults' Behavior and Knowledge in 2007. The Stationery Office, London.

Rehn, N., 2001. Alcohol in the European Region – Consumption, Harm and Policies. World Health Organization Regional Office for Europe.

Relevant Websites

http://www.ageuk.org.uk/ – Age UK.
http://www.alcoholconcern.org.uk/ – Alcohol Concern.
http://www.ifa-fiv.org/ – International Federation on Ageing.
http://www.niaaa.nih.gov/ – National Institute on Alcohol Abuse and Alcoholism.
http://www.se2009.eu/ –

CHAPTER

13

Relapse and Lapse

*Gillian Steckler**, *Katie Witkiewitz*[§], *G. Alan Marlatt*[†]

[*]Washington State University, Vancouver, WA, USA, [§]University of New Mexico, Albuquerque, NM, USA, [†]University of Washington, Seattle, WA, USA

OUTLINE

Many find that quitting a problem behavior is difficult and that maintaining the behavior change is even more of a challenge. The initial transgression of a problem behavior after a quit attempt, herein defined as a "lapse," could eventually lead to continued transgressions of the problem behavior to a level that is similar to the prequit level of behavior, which is often defined as a "relapse." For example, 40–80% of patients receiving treatment for alcohol use disorders have at least one drink, a "lapse," within the first year of after treatment, whereas fewer than 20% of patients return to pretreatment levels of alcohol use. Most psychological disorders and problem behaviors have very high rates of lapses, with many individuals having multiple "lapses" after treatment or a quit attempt. Thus, it is quite difficult to define the point at which "relapse" occurs and various definitions of relapse can result in practical differences in the interpretation of treatment effectiveness and patient prognosis. Based on the inconsistencies and difficulties in defining relapse, many researchers have proposed to define relapse as a process, rather than a discrete event.

Drawing from this conceptualization of relapse, one of the primary goals of problem behaviors research has been to characterize the relapse process using observable/measurable client characteristics as predictors or determinants of the relapse process. The purpose of this research is to gain a better understanding of predictors in the relapse process to implement relapse prevention. Relapse prevention (RP) is a psychotherapeutic approach that aims to teach clients skills in recognizing and avoiding high-risk situations for relapse through a cognitive-behavioral approach developed by Marlatt. Marlatt's cognitive-behavioral model of relapse prevention has been described as one of the most impressive and functional approaches in the field of addiction treatments.

The goal of this article is to review research on those factors that may predict the relapse process, as well as a review of the relapse prevention approach. We will

[†]Deceased

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00013-9

also suggest future clinical treatments and research questions pertaining to relapse prevention.

MODELS OF RELAPSE

The majority of people who have gone through treatment for substance addiction will experience some kind of "triggering" event, leading to an initial fall back to the undesired behavior (e.g. substance use), known as a lapse. Research has indicated that the majority of individuals who receive treatment for a substance use disorder will experience a lapse. Among those people who lapse, many are able to follow this with a "prolapse," defined as getting back on track in the direction of positive behavior change. This can be followed by remission, a relatively brief period of time without indulging in the behavior, which is often followed by another lapse cycle.

Cognitive-Behavioral Model of Relapse

Marlatt's cognitive-behavioral model of relapse has been an influential theory of relapse to addictive behaviors. The model defines the relapse process as a progression centered on "triggering" events, both internal and external, that can leave an individual in high-risk situations and the individual's ability to respond to these situations. In this process, after experiencing a trigger, an individual will make a series of choices and thoughts that will lead to being placed in a high-risk situation or not. There are two major types of high-risk situations, those with intrapersonal determinants, in which the person's response is physical or psychological in nature, and interpersonal determinants, those that are influenced by other individuals or social networks.

Starting from the point of confronting and recognizing a high-risk situation, Marlatt's model illustrates that the individual will deal with the situation with either an effective or ineffective coping response. Effective coping skills can lead to increased self-efficacy, and a decreased probability of a lapse. However, if one lacks skills, then the model predicts a decrease in self-efficacy and an increase in positive outcome expectancies for the effects of using the substance. This is a likely predecessor of giving into temptation in the initial use of a substance.

If an individual uses a substance after experiencing a remission, he/she may be vulnerable to the abstinence violation effect (AVE), which refers to an individual's response to the recognition that he/she has broken a self-imposed rule by engaging in substance use or other unwanted behavior. This response often creates a feeling of self-blame and loss of perceived control due to breaking a self-imposed rule regarding substance use. According to AVE research, those who do chose to respond to their behavior with blame and a sense of lost perceived control are more likely to relapse than those who respond by attributing lapse to preventable events and not feeling as though they failed completely. So long as an individual maintains a perceived sense of self-control, he/she has a better chance at evading further lapses. AVE has been studied and supported for the cessation of sex offenses, heroin, marijuana, and other illicit drug use.

Despite the empirical support for many components of the cognitive-behavioral model, there have also been many criticisms of the model for being too static and hierarchical. In response to these criticisms, Witkiewitz and Marlatt proposed a revision of the cognitive-behavioral model of relapse that incorporated both static and dynamic factors that are believed to be influential in the relapse process. The "dynamic model of relapse" builds on several previous studies of relapse risk factors by incorporating the characterization of distal and proximal risk factors. Distal risks, which are thought to increase the probability of relapse, include background variables (e.g. severity of alcohol dependence) and relatively stable pretreatment characteristics (e.g. expectancies). Proximal risks actualize, or complete, the distal predispositions and include transient lapse precipitants (e.g. stressful situations) and dynamic individual characteristics (e.g. negative affect, self-efficacy). Combinations of precipitating and predisposing risk factors are innumerable for any particular individual and may create a complex system in which the probability of relapse is greatly increased.

The revised dynamic model of relapse also takes into account the timing and interrelatedness of risk factors, as well as provides for feedback between lower- and higher-level components of the model. For example, based on the dynamic model it is hypothesized that changes in one risk factor (e.g. negative affect) influences changes in drinking behavior and that changes in drinking also influences changes in the risk factors. The dynamic model of relapse has generated enthusiasm among researchers and clinicians who have observed these processes in their data and their clients.

DETERMINANTS OF RELAPSE

Self-Efficacy

Self-efficacy is an important intrapersonal determinant of relapse and has been shown to be a predictor of outcomes across all types of addictive behaviors, including alcohol use, gambling, smoking, and drug use. Although essential in relapse prevention, measuring

self-efficacy is a challenge, since assessing it is limited within constrained contexts, rather than having the ability to measure it in individual high-risk situations. It is also difficult to use measures of self-efficacy as predictors of relapse because self-efficacy has been shown to be very highly correlated with other intrapersonal determinants of relapse.

When one group of researchers monitored individuals who had been diagnosed with substance dependence and a major depressive disorder, a significant association was established between self-efficacy, life stress, and periods of time to lapse for those who lapsed during the first 6 months in treatment. The study also found that individuals with high self-efficacy levels were able to maintain abstinence from drugs and alcohol for longer periods of time than those displaying lower levels of self-efficacy. Nearly every participant in the study who reported lower levels of self-efficacy experienced a lapse during the early stages of treatment. In addition to self-efficacy, the researchers found that acute stressors nearly tripled the risk of relapse, leading to poorer long-term outcomes for those in treatment.

Chronic stressors may also overlap between self-efficacy and other areas of intrapersonal determinants, like emotional states, by presenting more adaptational strain on the treatment-seeking patient. Chronic stressors are likely to result in depletion of coping skills and resources, which may reduce self-efficacy and leave the individual more vulnerable to high-risk situations. The combination of chronic stress, reduced self-efficacy, and negative emotional states greatly increases the probability of using a substance in a high-risk situation. Like the previously mentioned study, another group of researchers found that it is possible that a decrease of self-efficacy is likely to be influenced by inferior coping strategies, craving, negative affect, depletion of self-regulatory resources, and lack of knowing efficient coping skills. Specifically, these researchers found that the day before an initial smoking lapse, the level of individual self-efficacy plummeted to particularly low levels. At the same time that self-efficacy is falling, coping mechanisms to resist the urge to use may also drop, increasing negative affect and, in turn, increasing the risk of a relapse.

Outcome Expectancies

Outcome expectancies can be defined as an individual's anticipation or belief of the effects of a behavior or future experience. These expectancies greatly influence the behavioral response, depending on the strength and attractiveness of the expected outcome to the individual. Whether or not the substance in question is real, or even a placebo, the expectancy of the effects of the drug plays a major role in personal experience of a drug. Outcome expectancy theory proposes that individuals have expectancies about the drugs they use and that these expectancies are the result of both direct (e.g. memories of using at a party) and indirect (e.g. perceptions of a drug that has been derived from the media) experiences.

Outcome expectancies can be either positive or negative. Positive outcome expectancies, beliefs an individual holds that make a substance or behavior more attractive to the individuals (e.g. "That drink would make me feel relaxed"), are associated with poorer treatment outcomes and are an important component of motivation to drink. On the other hand, an individual's negative expectancies (e.g. "If I use heroin, I will feel awful later on") can motivate the individual to restrain, leading to improved outcomes and lesser risk of relapse.

According to Gwaltney, a clinical-health psychologist focused on smoking cessation, lapses are likely to occur when self-efficacy regarding abstaining is particularly low and when positive outcome expectancies for smoking are exceptionally high, with the influence of expectancies on lapse-behavior dependent on self-efficacy. Gwaltney and colleagues propose that the interconnectedness of outcome expectancies and self-efficacy can be attributed to the fact that they are both influenced by increases in negative affect or craving levels, in turn, increasing positive outcome expectancies and decreasing self-efficacy. While expectancies are related to outcomes, there is little evidence to support the notion that targeting outcome expectancies alone will change overall substance use or behaviors once the individual has left treatment.

Craving

Craving may be the most studied and most poorly understood concept in the field of addiction. Individuals with stronger cravings at admission into treatment are more likely to relapse. However, craving itself has not been shown to be a very strong predictor of lapse episodes. Interestingly, 78% of those who crave alcohol at admission to treatment do not show a diminished level of craving once discharged from treatment. It has been proposed that craving may not be a particularly strong predictor of lapse itself, but that the correlates of craving, including negative affect and decreased self-efficacy could be stronger predictors. Other factors that could be related to craving include perceived availability of a substance during a time of abstinence, symptoms of withdrawal, anticipated physical effects, and outcome expectancies of the drug, and fundamental conditioned responses to drug effects.

Numerous human and animal studies over the past 10 years have added more questions than answers to

the craving literature. For example, recent clinical studies have shown that craving for cocaine during treatment predicts post-treatment cocaine use and that cocaine use is highly correlated with cocaine craving in a person's natural environment, whereas other studies of cocaine craving have not found a significant association between craving and use. Similarly, one group of researchers found that cue-provoked craving predicted smoking lapses, whereas another researcher concluded that there was no compelling evidence to support the association between cue-induced craving and smoking lapses. Cross-substance craving could also be an important predictor of relapse as seen in a recent study which showed that smoking and craving for nicotine were associated with episodes of heroin and cocaine use, as well as craving for heroin and cocaine.

One recent method for managing craving and its correlates incorporates cognitive and behavioral coping skills through urge-surfing and mindfulness-based relapse prevention techniques. Researchers have found that mindfulness-based instructions are associated with significantly less alcohol and drug use after treatment, which is mediated by reductions in craving after the mindfulness-based intervention.

Motivation

Motivation may relate to the relapse process in two distinct ways, the motivation for positive behavior change and the motivation to engage in the problematic behavior. Using the example of alcohol use, we could define the first type of motivation, *motivation to change*, as the stimulus for action toward abstinence or reduced use of alcohol, and the second type of motivation, *motivation to use*, as the stimulus for engaging in drinking behavior. This distinction captures the ambivalence that is experienced by individuals attempting to change their addictive behavior. The ambivalence toward change is often highly related to both self-efficacy (e.g. "I really want to quit drinking, but I do not think that I'll be able to say no") and outcome expectancies (e.g. "I would quit drinking, but then I would have a really hard time meeting people"). With regard to the *motivation to change*, the trans-theoretical model of behavior change provides a framework for understanding motivation to change, defined by five stages of behavior change: pre-contemplation, contemplation, preparation, action, and maintenance. During pre-contemplation, there is very little motivation to change, but as the individual moves toward contemplation there is an increase in ambivalence and "change talk." Interventions that focus on resolving ambivalence (e.g. evaluating the pros and cons of change vs. no change) may increase intrinsic motivation by allowing the client to explore their own values and how they may differ from actual behavioral choices (e.g. "I want to be an effective employee, but I often spend my daytime hours hungover and my evening hours getting drunk."). One such intervention, called Motivational Interviewing (MI), was developed as a treatment technique for enhancing motivation and encouraging positive behavior change. Originally developed for patients presenting with alcohol use disorders, MI has demonstrated efficacy for reducing alcohol consumption and frequency of drinking in this population. Additionally, MI has been successfully adapted and applied to work with a variety of other health behaviors including weight reduction, use of illicit substances, smoking, and HIV risk reduction.

Motivation to use in any given situation is based on the reinforcement value of the outcome in that situation. Similar to outcome expectancies, craving, and self-efficacy, changes in motivational levels may serve as a marker or lapse risk preceding a lapse. This has been illustrated by observed abrupt drops in motivation levels in the weeks before relapse. Although research shows that these drops occur, little research has been conducted in developing measures that look at changes in motivation levels for smokers who have already attempted to quit smoking; most previous measures have been focused on motivation to quit smoking before the smoker has attempted.

Recent measures have been developed and are currently undergoing further analyses to measure abstinence-related motivational engagement (ARME). In previous studies, measures to assess motivation as a predictor in quitting a behavior and staying abstinent have generally been evaluated through single broad motivational items (e.g. on a Likert scale, "How motivated are you to quit smoking?"). Some researchers believe that this approach focuses too much on the desire to remain abstinent, rather than the motivation to put forth the necessary efforts it will require to remain abstinent. In contrast to the former, single-item measures of motivation to stay abstinent, the 16-item ARME scale reflects four themes of an individual's motivation to not use: cognitive effort, focusing, and thinking of staying abstinent; priority in comparison to other issues in the individual's life; vigilance and preparing for high relapse risk situations; and excitement and enthusiasm to remain abstinent. Although this information is new, initial studies for ARME to be used in targeting smoking relapse prevention therapies and predicting cessation results is promising.

Coping

The cognitive-behavioral model of relapse focuses on effective coping as central to RP. Coping skills have been characterized as either cognitive (e.g. urge-surfing) or behavioral (e.g. calling an abstinent friend) and as

either approach (e.g. calling a friend) or avoidance (e.g. avoiding a non-abstinent friend) coping. The behavioral approach has shown to be important in preventing alcohol lapses, while in another study, the cognitive approach was related to lower rates of heroin use relapse.

A 2004 study suggests that individuals with higher approach scores were more likely to be more engaged in treatment, were more willing to disclose lapses, and did as well at improving relapse knowledge as the avoidance-oriented group. The study also proposed that approach-oriented participants may see themselves as more responsible for their actions, including lapse, while avoidance-based coping may focus more on their environment than on their own actions, leading more favorably toward the approach coping method. In a 2007 study, avoidance coping strategies suggested a higher probability of alcohol use, specifically when there is a lack of social or familial support.

There is also support for avoidance-based coping strategies in RP, which are consistent with cognitive-behavioral treatments. While testing for the validity of a new measure for approach and avoidance coping, researchers utilized the Approach and Avoidance of Alcohol Questionnaire (AAAQ). The questionnaire was given to a group of alcohol-dependent individuals, results showing that higher approach scores could be correlated with heavier, more frequent, and more recent drinking, while higher avoidance scores were associated with fewer drinks in the past week and more time elapsed since their last drink.

There are arguments for approach- and avoidance-based coping in RP, but in general, coping skills and their different approaches are useful for changes in substance use, and can also be beneficial in management of stress, sex abuse, and craving. Importantly, it was also found that changes in coping during and after coping skills training programs (including RP-type interventions) did not mediate treatment outcomes. In other words, coping skills are related to outcomes, but changes in coping skills are not a powerful mechanism of treatment change after coping skills interventions.

Emotional States

Substance abuse relapse has been linked to negative affect in a number of studies and negative emotional states were identified as the most salient risk factor for relapse in Marlatt's early work in relapse prevention. Recently, researchers asked whether lapses result in an individual feeling more depressed or angry or anxious. Do changes in negative affect cue an individual to resume drinking? Or perhaps both processes are working concurrently.

Results of the questions posed in research yield support for a dynamic association between negative affect and alcohol use such that changes in alcohol use predicted subsequent changes in negative affect and changes in negative affect predicted subsequent changes in alcohol use. Results also suggest that reducing negative affect after alcohol treatment can increase an individual's chance of staying sober, or returning to sobriety after a brief lapse.

Placing significance on successful affect regulation to prevent relapse is a main theme in addiction treatment. For Alcoholics Anonymous (AA), individuals learn skills in coping, motivation, self-efficacy, social changes, and spirituality; however, anger is exclusively tagged as a high-risk emotion for relapse to alcohol. While examining anger as a predictor of relapse, researchers of a 2010 study found that anger levels at intake were found to be significantly higher than those of the general population. The mean score measuring anger for the group was in the 98th percentile, dropping to the 89th percentile 15 months preceding initial intake. It is worth noting that those who score above the 75th percentile are likely to experience anger intense enough to interfere with relationships, psychological and physical disorders, and psychosocial functioning. These factors stemming from a negative emotional state are prone to having a negative effect on self-efficacy, motivation, positive outcome expectancies, coping skills, and interpersonal determinants of lapse. Although emotional state is not a significant lone predictor of lapse, the findings suggest that treating depressive, negative affect symptoms after a lapse may be crucial to cessation.

Interpersonal Determinants

Interpersonal determinants include the social support system, level of perceived emotional support, and peer pressures to use substances. Without a proper support group, which should consist of high levels of high-quality support, an individual is likely to relapse. Positive, high-quality support networks are highly predictive of long-term abstinence rates. Sources of support are most commonly found in family, a spouse, friendships, and coworkers.

While many families can find addiction difficult to deal with, negative family behaviors such as withdrawing from a family member with a substance use disorder or avoiding dealing with the substance use is actually associated with more drinking. As challenging as it may be, good family functioning and adjustment are related to more positive substance use outcomes. This is also connected to spousal relationships. Although positive spousal behaviors are associated to more positive substance use outcomes, negative behaviors have a contrary effect. In a marital relationship, predictors for positive substance use outcomes include marital happiness, having a better-functioning marriage before

treatment, and being a married male. Predictors of negative outcomes include marital dissatisfaction, marital events, and negative spousal factors and behaviors.

In general, encouragement from friends and coworkers is also related to positive substance use outcomes. In addition to encouragement, other positive factors include surrounding oneself with more nondrinking friends and abstinent coworkers. On the other side of this, having more drinking or using friends, having just one friend who uses the same drug, and having a social network with higher stress levels are indicators of negative outcomes.

In addition to familial, marital, friend, and coworker relations, research also suggests that community-based support services, such as recovery communities, enhance outcomes for those struggling with cessation once treatment has been completed. In these peer-support communities, members of the group share their experiences and help one another cope with the issues, including minor lapses while maintaining a healthy, abstinent social network and enhancing their chances at recovery and abstinence. If a major lapse occurs, recovering peers are not expected to treat the individual; rather, he/she will be readmitted into treatment and invited back to the community when the level of recovery is appropriate.

Unfortunately, the relationship between substance treatment outcomes and interpersonal determinants is not clear-cut. Interpersonal relationships and support systems are highly influenced by intrapersonal processes such as emotion, coping, and expectancies, which is why some have suggested an extension of the Witkiewitz and Marlatt model of relapse which posits ways in which interpersonal behavior may influence their social network and vice versa.

RELAPSE PREVENTION

RP aims at educating the patient about the relapse process, likelihood of relapse, and how to provide ways to make it through attempts at giving up the problem behavior. Educating a treatment-seeking individual includes ongoing assessment of specific high-risk situations that may jeopardize an individual's attempt to maintain a treatment goal. These high-risk situations are identified through assessment of interpersonal, intrapersonal, and environmental factors, as an assessment of withdrawal and craving (physiological risks for relapse). Although these high-risk scenarios cannot always be planned for, educating the patient is aimed at helping him/her prepare for a high-risk situation through behavioral skills training and cognitive interventions aimed at limiting the occurrence of a lapse episode. Once high-risk situations are acknowledged, patients learn to identify and modify coping skills used, in turn, reinforcing self-efficacy. The application of new, superior coping skills greatly increases the likelihood of not only achieving, but also maintaining abstinence from the externalized behavior.

Effectiveness and Efficacy of Relapse Prevention

A meta-analysis of RP based on 26 controlled trials suggested RP was particularly effective for decreasing alcohol and multiple substance relapse, with somewhat weaker effects for smoking cessation. Relapse prevention is not only effective for alcohol and drug studies, it is also effective in the treatment of depression, sexual offenses, erectile dysfunction, obesity, eating disorders, panic disorders, obsessive compulsive disorder, schizophrenia, bipolar disorder, gambling, cocaine abuse, and marijuana use.

In these cases, RP has been proven to be more effective than no treatment at all, and equally effective as other treatments (e.g. supportive and interpersonal therapies). RP may also provide continued improvement over a longer period of time when compared with other treatments. In recent meta-analyses, the effectiveness of cognitive-behavioral RP therapy across studies were examined, while controlling for cross-study differences in target substance use, sex of the patient, format of treatment (group vs. individual), concurrent treatment, duration of the intervention, and long-term effects. Results of the study indicated cognitive-behavioral interventions (including RP) may be best suited for marijuana use disorders, with women, when provided in a brief format, and when used in combination with psychosocial treatments. These researchers have also found that the intervention proved equally effective when presented in a group or individual format.

PHARMACOTHERAPY APPROACHES TO RELAPSE PREVENTION

Nicotine-Replacement Therapies and Prescription Smoking Cessation Aids

Nicotine-replacement therapy (NRT) aids in the cessation of cigarette smoking by reducing physiological withdrawal symptoms (e.g. depression, trouble sleeping, irritability, anxiety) and craving, thus reducing motivation to smoke and increasing the likelihood of remaining abstinent. The amount of nicotine in NRTs is enough to increase levels of dopamine and norepinephrine, which are thought to provide the pleasurable effects of smoking cigarettes, as well as reduce withdrawal symptoms and craving. Prescription smoking cessation aids, including bupropion and varenicline,

act as dopamine reuptake inhibitors, associated with increases in extracellular dopamine and norepinephrine levels. These chemicals mimic nicotine reinforcement while alleviating symptoms of withdrawal, making nicotine withdrawal and craving easier to manage.

Cessation Aids for Alcohol Dependence

The most common cessation aids for alcohol dependence are acetaldehyde dehydrogenase (ALDH) inhibitors, and opiate antagonists. Disulfirum is an ALDH inhibitor that, when mixed with alcohol, causes a buildup of toxic acetaldehyde resulting in flushing, headache, nausea, vomiting, chest pain, weakness, blurred vision, sweating, choking, respiratory difficulty, as well as mental confusion and anxiety. It is primarily used because of its unappealing reaction when introduced to alcohol, but also increases dopamine concentrations causing some individuals to report a decrease in alcohol craving. Disulfirum has become a key component in preventing lapse and relapse for many patients due to alcohol avoidance. Other alcohol-cessation drugs, such as naltrexone, block opioids from the "rewarding" pathways of the brain, neutralizing positive responses, making drinking less gratifying, and making it easier to stop using alcohol. Due to fewer reward responses, craving is reduced and patients are less likely to relapse to heavy drinking. Acamprosate is the latest drug to be approved by the FDA for alcohol dependence treatment. Results of numerous studies suggest that acamprosate may function by relieving some of the withdrawal symptoms of alcohol dependence and lessening the likelihood of lapse and relapse. Results show consistent improvement in abstinence rates and duration of abstinence when combined with psychosocial interventions.

Illicit Drug Cessation Aids

The most successful pharmacotherapies for the prevention of illicit drug relapse have been opioid substitution treatments developed for opiate dependence, including methadone and buprenorphine. Opioid substitution treatments provide a safer alternative to heroin and other opioids (e.g. oxycodone or hydrocodone, which are legal and prescribed for pain, yet widely available on the street for recreational use). At lower, prescribed doses, cessation aids give addicted individuals the opportunity to abstain from opioids without experiencing the severe withdrawal symptoms that often lead one to lapse and relapse.

The knowledge of pharmacotherapy for addiction is constantly growing and many new cessation drugs are being investigated. Importantly, all of these pharmacotherapies have potential side effects and are not effective for everyone. Furthermore many studies have shown that many of the medications described above are most effective when combined with psychotherapy.

SEE ALSO

Behavioral Economic Factors in Addictive Processes, The Biopsychosocial Model of Addiction, Cognitive Factors in Addictive Processes, Stress and Addiction

Glossary

Abstinence violation effect (AVE) an individual's response to the recognition that he/she has broken a self-imposed rule by engaging in the problem behavior, creating a feeling of self-blame and loss of perceived control. Individuals who respond in this manner are more likely to relapse than those who attribute lapse to preventable events that are controllable.

Craving an intrapersonal determinant of relapse of addictive behaviors; wanting more of something (e.g. alcohol, a drug).

Lapse the initial transgression of a problem behavior after a quit attempt.

Relapse continued transgressions of a problem behavior to a level that is similar to the prequit level of behavior; a process, rather than a discrete event.

Self-efficacy the belief in one's capabilities to achieve a goal or an outcome or the belief in one's ability to succeed in specific situations; an intrapersonal determinant of relapse and shown predictor of outcomes across addictive behaviors.

Further Reading

Baker, T.B., Piper, M.E., McCarthy, D.E., Majeskie, M.R., Fiore, M.C., 2004. Addiction motivation reformulated: an affective processing model of negative reinforcement. Psychological Review 111, 33–51.

Bowen, S., Marlatt, A., 2009. Surfing the urge: brief mindfulness-based intervention for college student smokers. Psychology of Addictive Behaviors: Journal of the Society of Psychologists in Addictive Behaviors 23 (4), 666–671.

Connors, G.J., Tarbox, A.R., Faillace, L.A., 1993. Changes in alcohol expectancies and drinking behavior among treated problem drinkers. Journal of Studies on Alcohol 54, 676–683.

Donovan, D.M., 1996. Marlatt's classification of relapse precipitants: is the Emperor still wearing clothes? Addiction 91, 131–137.

Drummond, D.C., Litten, R.Z., Lowman, C., Hunt, W.A., 2000. Craving research: future directions. Addiction 95, 47–55.

Edwards, G., Marshall, E.J., Cook, C.C., 1999. The Treatment of Drinking Problems: A Guide for the Helping Professions, third ed. Cambridge University Press, Cambridge.

Fiore, M.C., Hatsukami, D.K., Baker, T.B., 2002. Effective tobacco dependence treatment. JAMA 288, 1768–1771.

Gossop, M., Stewart, D., Browne, N., Marsden, J., 2002. Factors associated with abstinence, lapse or relapse to heroin use after residential treatment: protective effect of coping responses. Addiction 97, 1259–1267.

Gwaltney, C.J., Shiffman, S., Balabanis, M.H., Paty, J.A., 2005. Dynamic self-efficacy and outcome expectancies: prediction of smoking lapse and relapse. Journal of Abnormal Psychology 144, 661–675.

Hunt, R.R., 2002. Clinical inquires. How effective are pharmacologic agents for alcoholism? The Journal of Family Practice 51, 577.

Hunter-Reel, D., McCrady, B., Hildebrandt, T., 2009. Emphasizing interpersonal factors: an extension of the Witkiewitz and Marlatt relapse model. Addiction 104, 1281–1290.

Magill, M., Ray, L.A., 2009. Cognitive-behavioral treatment with adult alcohol and illicit drug users: a meta-analysis of randomized controlled trials. Journal of Studies on Alcohol and Drugs 70, 516–527.

Marlatt, G.A., 1985. Relapse prevention: theoretical rationale and overview of the model. In: Marlatt, G.A., Gordon, J.R. (Eds.), Relapse Prevention. Guilford Press, New York, pp. 250–280.
Witkiewitz, K., Marlatt, G.A., 2004. Relapse prevention for alcohol and drug problems: that was then, this is Tao. American Psychologist 59, 224–235.
Witkiewitz, K., Villarroel, N.A., 2009. Dynamic association between negative affect and alcohol lapses following alcohol treatment. Journal of Consulting and Clinical Psychology 77, 633–644.

Relevant Websites

http://nationalpsychologist.com/ – Relapse Prevention Therapy.
http://www.addictioninfo.org/ – News and Views on Substance Abuse.
http://www.samhsa.gov/ – Substance Abuse and Mental Health Services Administration.
http://dailyuw.com/2007/5/2/whats-new-in-science/ – What's New in Science.
http://www.mindfulness.org.au/ – Mindfulness.

CHAPTER

14

Natural Recovery

Tony Toneatto

University of Toronto, Toronto, ON, Canada

OUTLINE

INTRODUCTION

Traditional conceptual models of addictive behavior have stressed their chronic, progressive, and unremitting nature. Among many expressions of this perspective, the disease model of addiction, whether reflected in the self-help emphasis on the spiritual aspects of addictive disease, or the more recent, neuroscientific view of addiction as a "brain disease," the assumption is that addiction is mediated by basic deficiencies in brain, mind, or spirit. In the absence of treatment, and left to its own natural course, death, disability, and deterioration are the inevitable outcomes. Without the active intervention of professionals, within the context of formal treatment, the assumption is that a durable recovery would be highly unlikely. Such views of addiction and its resolution have also generally been held by the general public. This model of addictive behavior and recovery, based on the study of individuals with severe dependence or who have been in formal treatment programs, and an emphasis on lifelong abstinence, leaves little space for alternative approaches and outcomes. As a result, a bias arguing against the possibility of natural recovery from addiction has resulted. This chapter will review the recent empirical literature describing what is known about individuals who recover from addiction without formal psychological treatment. Although the majority of the research has focused on the abuse or dependence on alcohol, studies of natural recovery among other psychoactive dependencies and pathological gambling will also be cited.

The assumptions of the disease model have been seriously challenged in the past several decades with an accumulating body of empirical research demonstrating that recovery from addiction may commonly occur in the absence of formal treatment and that such recoveries do not always lead to abstinence. Viewing addiction problems on a continuum of severity ranging between mild and severe dependence opens the possibility that there may be multiple approaches to resolution of addiction and goals other than total abstinence.

Natural recovery (or spontaneous remission, natural remission, untreated remission, spontaneous recovery, maturing out, self-change, auto-remission), the notion that the resolution of a serious addictive problem could occur in the absence of professional treatment and an adherence to abstinence, is one of the "taboo" topics in addiction treatment. Indeed, the presence of an addiction could only be verified if the individuals could not stop on through their own efforts. If one could cease an

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00014-0

addictive behavior, then it is unlikely that they were addicted in the first place. Such tautological reasoning served as "blinders" to the empirical literature by demonstrating the successful recovery by individuals who would be considered to be dependent on psychoactive substances by any diagnostic criteria. Although reports of recovery that did not conform to the received opinion had been available in the research literature, these did not significantly alter the basic view that addiction was a chronic, progressive disease that would ultimately lead to destruction and death unless remediated through professionally mediated interventions. The conviction with which this attitude has been held by generations of clinicians and researchers is perplexing when it is considered that the vast majority (over 80%) of individuals with nicotine addiction, for example, have recovered without formal treatment and that the rates of lifetime addiction are not matched by similar rates of lifetime treatment seeking.

The natural recovery research has also challenged the conventional notion that the control of the addictive process lies largely outside the volitional control of the individual. Overcoming addiction, within the traditional perspective, requires that the individuals surrender personal control over the addictive process to the influence and variables beyond them (e.g. a medication, Higher Power, God, therapist, physician, group process) and that the individuals' control lies primarily in the insight that such surrender is necessary. Without the intervention and assistance of a transcendental influence, the individuals cannot make therapeutic progress and durable recovery is thus unlikely. This assumption has been strongly challenged by the extensive natural remission and problem-solving therapy (i.e. cognitive-behavioral therapy) literature which has unequivocally and empirically demonstrated that individuals with substance dependencies do remain able to choose among a range of alternative resolution strategies (e.g. 12 Step programs, medication, residential treatment, outpatient therapy) including recovery without treatment even in the midst of an addictive process. While the severity of the addiction may make it less likely for an individual to choose natural recovery, such a choice remains a possibility and is often the first choice. It has been argued that formal addiction treatment is only sought if efforts to stop on one's own are unsuccessful.

In summary, although early studies of untreated recovery were controversial for their conflict with the traditional concept of addictive processes, the consensus today suggests that not only may natural recovery be the most common pathway to the resolution of an addiction, but that this process occurs across all addictive behaviors, substance-based or behavioral (e.g. pathological gambling). Self-recovery has been demonstrated for several psychoactive substances, including alcohol, cigarettes, cannabis abuse and dependence, cocaine, and other drugs. Individuals with behavioral addictions, such as problem gambling, have also been shown to recover without formal treatment. The Institute of Medicine (1990) has identified and supported natural recovery as an empirically supported route to recovery.

VALUE OF NATURAL RECOVERY RESEARCH

Apart from the scientific interest in studying untreated remission as a valid pathway to addiction recovery there are several clinically relevant reasons for studying this phenomenon. Firstly, what constitutes truly effective treatment is yet to be defined. Formal, professional treatment continues to evolve and no treatment is reliably and durably effective. The results of Project MATCH have demonstrated how difficult it is to successfully match treatments to individuals. In the absence of a therapeutic "cure" for addiction it behooves the clinical addiction research community to continue to deepen our understanding of the variables that lead to the resolution of an addiction problem. Furthermore, treatment seeking occurs for only a small proportion of all those who develop a clinically significant addiction. Indeed, the vast majority of addicted individuals never seek treatment. Recent comprehensive reviews of natural recovery have shown that addicted individuals may avoid treatment because they fear the associated stigma (a variable that may help account for the reluctance of some minority groups and women to seek treatment), believe their problem is not sufficiently severe to merit professional treatment, or have a preference to resolve their addiction problem on their own, relying on their personal resources. Thus, knowledge of the processes that mediate this increasingly common pathway to addiction recovery may improve interventions for individuals who wish to resolve without professional treatment.

In addition to the necessity of continuing to understand recovery, regardless of how it occurs, and of deriving a coherent scientific theory of untreated recovery, the elucidation of the variables that facilitate and maintain such behavior change may have important implications for the delivery of professional treatment, prevention strategies, and public policy. Clinicians can integrate the insights derived from the natural recovery literature into their formal interventions; prevention and awareness strategies can be developed based on the change processes that have been shown to trigger natural recovery; public policy can make this information available to a wider population of addicted individuals through a variety of means (e.g. web-based).

RATES OF NATURAL RECOVERY

Several studies in recent decades have repeatedly demonstrated that significant numbers of individuals with substance dependencies recover without ever receiving formal treatment. That such recoveries occur is no longer surprising or controversial. Epidemiological surveys of American adults found that almost three-quarters of those with alcohol dependence who had remitted in the previous year had done so without receiving professional help. These rates were similar to the rates among older adults (aged 51–65). Other research has found that as many as one out of five college students with a history of adolescent binge significantly reduced their alcohol consumption of drinking while still in college and without treatment.

The duration of the recoveries reported in these studies were neither transient nor unstable. An average of about 6 years has been reported across different psychoactive substances indicating that these recoveries are durable and sustained. Since a period of about 5 years is considered a stable recovery, with or without treatment, the length of recovery reported in spontaneous recovery studies suggest that these recoveries are not a mere respite from an ongoing addictive process. However, it should be noted that some studies have shown that the recovery from the primary substance does always mean that the individual is abstinent from all psychoactive substances and may thus be associated with the continued use or abuse of another substance (e.g. of alcohol in a sample of naturally recovered cocaine and heroin users).

REVIEW OF THE NATURAL RECOVERY LITERATURE

In their seminal review of 38 natural recovery studies (comprising 40 samples) dating back several decades, from 1960s to 1998, Sobell, Ellingstad, and Sobell (2000) found that the methodological quality of the research was generally poor. In general, early studies failed to describe their samples adequately, did not describe pre-recovery substance histories and severity in sufficient detail, failed to assess factors associated with the maintenance of recovery, ignored the complicating factor of concurrent psychiatric psychopathology, did not corroborate subject self-report, or assess family history of addiction. The vast majority (75%) of earlier studies addressed the spontaneous recovery from alcohol (not unusual since this substance is the most abused in western culture), followed from heroin (22.5%) and cocaine (7.5%). Advertisements (38.5%) were the most common way these subjects were recruited, followed by snowballing techniques (28%) and surveys (23%). The samples surveyed were generally male (69%), unmarried (64%), and unemployed (45%) with a mean age of 34.4 years at the time of their recovery.

Recently, Carballo et al. (2007) updated the Sobell et al. (2000) review. Twenty-two natural recovery studies were published in the 7 years following Sobell et al. (2000). Similar to the earlier review, the Carballo et al. study (2007) found that the majority of the 22 studies addressed natural remission from alcohol (82%) followed by cannabis (32%), heroin (23%), and cocaine (23%). Surveys (45%) and media solicitation (41%) remained the most common ways to recruit subjects. The demographic profile across both reviews was quite similar. Men continued to be the majority of individuals recruited into natural recovery studies, with a mean age in the early 40s, likely reflecting the higher prevalence of addiction among men. The Carballo et al. (2007) review found that the length of the addiction prior to natural recovery was approximately 13 years (compared to 11 years in the Sobell, et al. (2000) review). The mean length of the natural recovery was 8 years in the Carballo et al. (2007) review compared to 6.3 years in the earlier review. Abstinent outcomes continued to be reported by over half the sample in both reviews (57% versus 60%, respectively). Family-related reasons (54.5% of studies) with health (50%) followed by financial concerns (50%) were the most common explanations for the decision to recover from the addiction in the Carballo et al. (2007) review while health-related reasons were cited as the most important factor in the older review. The most common factors maintaining natural recovery reported in the Carballo et al. (2007) were social (54.5% of respondents) and family support (45.5% of respondents), similar to the 2000 review. About a third of the respondents reported avoidance of the addictive substance, self-control, and spirituality as important maintenance factors.

In their assessment of the quality of the natural recovery research since 2000, Carballo et al. (2007) noted that socio-demographic variables at the time of the recovery continued to be underreported (compared to these variables at the time of the interview, which generally tended to be well-reported). In other words, a clear understanding of the sample at the time of the recovery was generally lacking. Insufficient research examining the relationship between natural recovery from multiple substances has been conducted and remains a neglected area of research. Naturally recovered individuals tended to have had a less severe addiction with less severe consequences than did treated populations, a finding that has been repeatedly substantiated empirically.

The study of natural recovery from nonalcohol substances remain in the minority. The association between natural recovery and psychiatric comorbidity remains unclear although in a recent study alcohol natural recovery was equally prevalent among nonpsychiatrically comorbid or comorbid individuals suggesting that the presence of Axis I disorders was not a negative prognosis for untreated recovery from alcohol dependence. However, this should not be interpreted to mean that all natural recovery samples are similar. For example, three subgroups of naturally recovered populations have been delineated: low dependence/few alcohol problems/low social support; high dependence/several alcohol problems/moderate social support; low dependence/high social support/few alcohol problems. Such complexity precludes any simple description of naturally recovered populations.

PROCESS OF NATURAL RECOVERY

Having established the frequent occurrence of natural recovery, an issue of great interest is the trigger or reason for such transformations and the variables that may maintain such recovery. There appears to be a growing consensus that a cognitive process mediates natural recovery. In response to the debilitating effects of addictive behavior on their physical and mental health, relationship with significant others, and financial status, individuals undergo a process of rational re-evaluation of their substance use, weighing the advantages and disadvantages of continued use. Over time, the outweighing of the positives of substance use through the accumulation of negative consequences leads to a reconsideration of the benefits, value, and advantages of continued substance use. When the decisional balance tips irrevocably toward aversive effects, the individual is more likely to decide to make efforts to cease use. There are several models of cognitive change that are applicable to this decision-making process such as the Trans-theoretical Model of Change and Conflict Theory. However, little is known about what contributes to the defining decisional moment that leads to stable behavioral change.

This cognitive process mediating natural recovery is to be distinguished from recoveries triggered by the occurrence of discrete, highly salient, negative consequences, such as a stressful life event (e.g. severe illness, loss of relationship, major injury), that leads to a rapid decision to cease use. While these types of events certainly play a role, cognitive reappraisals appear to be the more common process and have been observed in the recovery from cocaine, heroin, and alcohol addiction as well as behavioral addictions such as problem gambling. In a qualitative analysis of recovery narratives among alcohol, cocaine, and heroin subjects, the following 11 major categories of reasons for recovery were identified: cognitive evaluations/assessments, behavior monitoring/action statements, problem-related reasoning, dramatic events, references to others/statements of support, health, religious attributions, time-frame, affect-related statements, alcohol-related statements, references to illicit/licit drugs. The most common reasons allude to an ongoing examination of the effects of the addictive behavior on the individuals' functioning. Evidence for a role of cognitive re-evaluation (versus discrete events) has implications for community-level interventions in which evaluation of the pros and cons of substance use is facilitated as a means of shifting the individual's attitude toward the substance in the direction of the cessation or reduction.

MAINTENANCE OF NATURAL RECOVERY

The maintenance of natural recoveries has received less empirical study. The few studies that have investigated this aspect of recovery have identified the critical role of social support from significant others. This finding replicates what is already known about the maintenance of recovery among treated samples. In the case of natural recovery from drug abuse, there is also a "geographic" change in which the individual avoids social and physical environments associated with drug use. The role of social capital, defined as the resources, available to an individual as a result of the availability of a network of mutual relationships, in the maintenance of untreated recovery can provide sources of social control, family support, and extra-familial benefits. Taking into consideration the individual's social capital necessarily directs our attention to the broader social context of recovery and the role that structured social relations play in overcoming drug-use related problems without treatment. How nonaddicts respond to the individual recovery and whether such responses can facilitate the recovery requires additional study. In a study of naturally recovered opiate addicts, the importance of social and familial relationships, gratifying leisure activities (e.g. hobbies), and employment in the maintenance of behavioral change were identified as critical to the maintenance of recovery. A conventional lifestyle that brings structure to one's life was considered crucial in establishing a successful recovery. It has been suggested that natural remission from drug dependence, in contrast to alcohol dependence, often requires a radical modification of their social network and way of life. Naturally

recovered alcohol abusers are more likely to return to a previously existing conventional lifestyle.

METHODOLOGICAL ISSUES

Natural recovery studies have been divided into two major types: those that investigate the variables associated with the onset and maintenance of natural remission (and which may include control groups) and those that estimate the prevalence of natural remission. The two classes of studies are complementary as the latter research provides the estimate of how widespread natural recovery is while the former elucidates the psychosocial variables that contribute to natural recovery. In recent years, general population studies have reduced their reliance on solicitation techniques based on snowballing or media thereby reducing the biases of self-selection. Studies that have relied on media recruitment may recruit individuals who are different than those who do not respond to media solicitations (e.g. less dependent). The variability of what is considered an alcohol problem (e.g. dependence, abuse, self-report) (not described in 40% of the studies reviewed by Sobell et al. (2000)) remains a serious weakness of the earlier literature. This variable is an important descriptive variable since, for example, natural remission may be correlated with the severity of the alcohol problem. This was demonstrated in a study where those with only one alcohol-related problem had elevated rates of natural remission (87.5%) compared to those who reported six alcohol-related problems (53.7%).

The definition of what constitutes treatment has varied across studies. This issue has been further complicated by the fact that the benefits of formal treatment may not always be correlated with the number of sessions attended. There is strong evidence that even very brief professional contacts, as may often occur in the offices of physicians, can have a significant impact on patients' behavior. In some studies of natural recovery, some subjects have even reported receiving formal treatment but have adamantly maintained that these treatments did not affect their recovery. In other cases, especially when natural recovery over the life span has been examined, formal treatment may have occurred years prior to recovery. Thus, studies that allow very minimal treatment contact, including self-help groups (i.e. two meetings), may describe a different sample of natural recovery than do studies that rigidly exclude any treatment contact whatsoever, including attendance at self-help groups. Many studies unfortunately combine individuals with a treatment history with those without such a history. In recent years, a stricter, more conservative definition of treatment has been employed and data separated for those with a positive history from those with no history. Naturally recovered samples with a history of formal treatment are not equivalent to those with no, or very minimal, histories of professional treatment. Typically, those with a history of treatment have more severe addiction problems than those without any treatment history.

NATURAL RECOVERY FROM PROBLEM GAMBLING

Although identified as an impulse disorder in DSM-IV the key symptoms defining this disorder and the recommended treatment approaches bear a very strong resemblance to substance use disorders. Not uncommonly, gambling is considered an addictive disorder among the professional community and the general public. Since the prevalence data suggest that the population of individuals who have a mild to moderate gambling problem may be much larger than those with a more severe problem, natural recovery as a common and preferred pathway to recovery requires additional study.

There have been a number of studies that have investigated the natural recovery from pathological or problem gambling. Not surprisingly, natural recovery from problem gambling appears to be highly comparable to that observed for substance use disorders. Several sources of evidence suggest that recovery from a gambling problem may not always be mediated through contact with the formal treatment system but rather reflect natural recovery processes paralleling those found with untreated recovery from substance dependencies. As has also been the case with substance use disorders, epidemiological studies of gambling prevalence frequently identify significant numbers of "former gamblers." For example, over a third of lifetime gamblers surveyed reported no problems in the previous year but did not report having received treatment. Similar results have been reported in American and Australian population surveys. In addition to the epidemiological evidence, the number of treatment-seeking gamblers is often considerably below what would be expected based on the point prevalence data. The National Gambling Impact Study Commission (1999) has estimated that less than 3% of pathological gambling had participated in formal treatment. Clearly, a substantial number of problem gamblers do not resolve this impulse disorder through formal treatment.

Similar to untreated recovery from addictions, natural recovery from problem gambling appears to involve a cognitive evaluation process focused on the detrimental impact of gambling on the individuals' core values as well as an accumulation of gambling-related negative consequences. Negative emotional

states, financial crisis, interpersonal distress, and conflict have been identified as frequently mentioned reasons for resolving the gambling problem in several studies. Change strategies reported by naturally recovered problem gamblers included staying from social situations and environments where powerful gambling triggers may be present, avoiding gambling locations, instituting desirable lifestyle change, and remaining aware of gambling-related negative consequences. Stimulus control, development of gambling-incompatible lifestyle, reduced access to money, and awareness of gambling-related negative consequences have been identified in many studies as the most common change maintenance strategies following recovery. The majority of naturally recovered gamblers, unlike what has been reported in the substance use literature, choose abstinence. Untreated, naturally recovered gamblers appear to have a less severe gambling problem than do treated gamblers. Gambling severity was found to predict treatment entry with less severe problem gamblers more likely to prefer natural recovery. Treated gamblers were found to have had a longer problem gambling duration, greater gambling severity, more numerous symptoms of gambling (e.g. feelings of despair, panic, suicide), and more gambling-related negative consequences (e.g. family, health) compared to the untreated but recovered gamblers.

In summary, the severity of problem gambling may be the main variable distinguishing those who choose to recover from gambling without treatment. Most studies of gambling natural recovery, like those of substance dependence, identify a crisis in self-image or values as accompanied by multiple gambling-related negative consequences as precipitating a re-evaluation process of the role of gambling in their lives.

CONCLUSIONS AND FUTURE DIRECTIONS

Untreated recovery from addictive behavior has been demonstrated to be a common pathway across all addictions. The results of a diverse set of methodological approaches have shown this recovery process to be robust. Several approaches to identifying naturally recovered individuals (e.g. population surveys, media solicitation, surveys, snowballing techniques) have shown substantial rates of recovery from alcohol and other drug problems without ever receiving formal treatment or attending self-help groups. Such recoveries have been shown to generally be triggered by cognitive appraisals of the cost and benefits of substance use. That is, addicted individuals reach a point in their addiction career in which they can no longer continue or justify addictive behavior. While specific events may be associated with the final decision to quit or reduce significantly, it is clear that a parallel cognitive process has been active for which the triggering event has been the definitive one. Such cognitive processes are likely present in all individuals caught in the addictive cycle and can ultimately lead to recovery as the benefits or advantages of substance use outweigh the decided disadvantages and risks.

Most evidence suggests that naturally recovered individuals are less severely dependent than those who recover with formal assistance. A review of the natural recovery research literature came to the following conclusions regarding the study of natural recovery (and applicable to the study of natural recovery from problem gambling as well). It was suggested that future research should focus on the following: ethnic subgroups (e.g. Asian), high-risk groups (e.g. aboriginals) and other minorities, individual differences characterizing those who recover without treatment, and gender differences. Major reviews of the natural recovery literature have also called for the study of subgroups within naturally recovered samples acknowledging that this population is not heterogeneous. Thus, future research should continue to explore individual differences in natural recovery in order to determine which subpopulations of substance dependence are likely to adopt natural recovery. Such information will be critical in developing targeted prevention and public health programs to reduce the incidence of substance abuse and dependence.

SEE ALSO

Alcohol Use Disorders, Gambling, Epidemiology of Addiction, Maturing Out

Further Reading

Biernacki, P., 1986. Pathways from Heroin Addiction Recovery Without Treatment. Temple University Press, New York.

Bischof, G., Rumpf, H.J., Hapke, U., Meyer, C., John, U., 2005. Natural recovery from alcohol dependence: how restrictive should our definition of treatment be? Journal of Studies on Alcohol 63, 229–236.

Carballo, J.L., Fernández-Hermida, J.R., Secades-Villa, R., Sobell, L.C., Dum, M., García-Rodríguez, O., 2007. Natural recovery from alcohol and drug problems: a methodological review of the literature from 1999 through 2005. In: Klingemann, H., Sobell, L.C. (Eds.), Promoting Self-change from Addictive Behaviors: Practical Implications for Policy, Prevention, and Treatment. Springer, New York, pp. 87–101.

Institute of Medicine, 1990. Broadening the Basis of Treatment for Alcohol Problems. National Academy Press, Washington, DC.

Klingemann, H., Sobell, L.C. (Eds.), 2007. Promoting Self-change from Addictive Behaviors: Practical Implications for Policy, Prevention, and Treatment. Springer, New York.

Klingemann, H., Sobell, M.B., Sobell, L.C., 2009. Continuities and changes in self-change research. Addiction 105, 1510–1518.

National Gambling Impact Study Commission, 1999. Final Report. U.S. Government Printing Office, Washington, DC.

Rumpf, H.J., Bischof, G., Hapke, U., Meyer, C., John, U., 2009. Remission from alcohol dependence without formal help: current status of the research. Sucht 55, 75–85.

Slutske, W.S., 2006. Natural recovery and treatment-seeking in pathological gambling: results of two U.S. National Surveys. American Journal of Psychiatry 163, 297–302.

Sobell, L.C., Ellingstad, T.P., Sobell, M.B., 2000. Natural recovery from alcohol and drug problems: methodological review of the research with suggestions for future directions. Addiction 95, 749–764.

Toneatto, T., Millar, G., 2004. The assessment and treatment of problem gambling: empirical status and promising trends. Canadian Journal of Psychiatry 49, 173–181.

Toneatto, T., Sobell, L.C., Sobell, M.B., Rubel, E., 1999. Natural recovery from cocaine dependence. Psychology of Addictive Behavior 13, 259–268.

CHAPTER

15

Gender Differences

Susan Nolen-Hoeksema
Yale University, New Haven, CT, USA

OUTLINE

GENDER DIFFERENCES IN ALCOHOL USE AND RELATED PROBLEMS

Dozens of studies done across many countries and cultures of the world have established that men are more likely to drink alcohol than women, men drink a greater volume of alcohol than women if they drink at all, and men are more likely to show behavioral problems as a result of alcohol use than women. For example, analyses of data from 35 countries by the GENACIS project (Gender, Alcohol, and Culture: An International Study) show that drinking and high-volume drinking were consistently more prevalent among men than among women, whereas lifetime abstention from alcohol was consistently more prevalent among women. Among drinkers, women in all age groups were consistently more likely than men to have stopped drinking. The size of the gender ratios, and the absolute amount of drinking, varied considerably across country. For example, rates of heavy episodic drinking ranged between 91 and 62% of male and female drinkers, respectively, in Ireland and to 22 and 11% of male and female drinkers in Israel. Still, men evidenced more drinking and related problems than women in all age groups in almost all countries.

Gender differences in alcohol misuse disorders are also consistently found. In the United States, the National Comorbidity Survey (NCS), which surveyed over 8000 adults under 55 years of age, estimated that, at some time in their lives, 12.5% of men and 6.4% of women will meet the criteria for alcohol abuse, and that 20.1% of men and 8.2% of women will meet the full criteria for alcohol dependence. Similarly, the National Longitudinal Alcoholic Epidemiologic Survey (NLAES), which had a nationally representative sample in the United States of 42 862 adults 18 years of age and older, and used DSM-IV criteria to diagnose alcohol dependence, found a lifetime prevalence for alcohol dependence of 18.6% for men and 8.4% for women. Although most cross-national studies have examined gender differences in problems related to drinking rather than diagnosed alcohol misuse disorders per se,

http://dx.doi.org/10.1016/B978-0-12-398336-7.00015-2

in those that have examined diagnoses, males are more consistently more likely than females to meet criteria.

Some studies suggest that, in recent decades, the gender difference in drinking and alcohol-related problems has decreased. For example, retrospective data from the NCS suggest that there has been a convergence in recent decades between men and women in their probabilities of alcohol use and problems, with both genders showing earlier onset of drinking and symptoms of alcohol use disorders in more recent cohorts than in cohorts born a few decades ago. Cross-sectional data from surveys of high school students in the United States from 1975 to 2001 suggest that the gender gap in the prevalence of binge drinking (five or more drinks in a row) decreased from 22.6% in 1975 to 12.3% in 2001. The typical explanation for this convergence is that females are drinking more and having more related problems in recent years, due to decreases in social pressures against drinking by females. Analyses of data from a longitudinal study of women from 1981 to 2001 provide mixed support for this explanation, at best. Researchers found increases over time in the percentage of women who drank at all. There were also increases over time, however, in abstinence and declines in heavy episodic drinking among women. Moreover, although the size of the gender gap in alcohol use and related disorders may be decreasing, women still consume significantly less alcohol and are less likely to manifest alcohol-related problems than men.

The comorbidity between alcohol misuse disorders and diagnoses of depression, anxiety disorders, eating disorders, and borderline personality disorder is higher for women than for men. The comorbidity between alcohol misuse disorders and antisocial personality disorder is greater for men than for women.

SOCIAL CONTEXT FOR MEN'S AND WOMEN'S DRINKING

For both men and women, drinking is often a social activity, but the social context for drinking differs somewhat for men and women. For men, drinking often occurs in the context of peer or work relationships (i.e. going to a bar with buddies or coworkers to relax and have a few). This social activity appears to strengthen these relationships and serve as a means of social support. Although women also drink with friends and coworkers, heterosexual women's levels of drinking are particularly strongly tied to husbands' or male partners' levels of drinking. Women problem drinkers are more likely to be married to male problem drinkers than male problem drinkers are to be married to women problem drinkers. This may be due to the fact that there are more male problem drinkers than there are women problem drinkers. But when a husband's and wife's drinking patterns are discrepant, there is more general discord in the marriage. A movement toward more similar drinking patterns in a couple may be reinforced by the resolution of marital discord. A fascinating twin study provided evidence that marriage does affect how women drink: Andrew Heath and colleagues found that genetic factors were more strongly related to women's drinking patterns if they were unmarried than if they were married, suggesting that marriage exerted a significant effect on women's drinking patterns over and above the effect of their genetic heritage.

BIOLOGICAL CONTRIBUTORS TO THE GENDER DIFFERENCE

The universality of the gender difference in alcohol use and problems suggests that biological factors may play a role. Differences in the effects of alcohol on intoxication in women and men, and genetic factors, have been argued to contribute to the gender difference in alcohol and related problems.

Effects of Alcohol in Men and Women

After an equivalent dose of alcohol, women have higher blood ethanol levels than men. This may lead women to drink less and thus to show fewer alcohol-related behavioral problems and be less likely to develop tolerance to alcohol than men. Women's higher blood ethanol levels at a given dose may occur because women are generally smaller than men, or because women's body water content is smaller than men's per kilogram of body weight, leading to greater amounts of alcohol being passed through the stomach and into the blood stream in women compared to men. Also, gender differences in the metabolism of alcohol may lead to gender differences in blood alcohol concentrations following a given dose. In the stomach, alcohol is metabolized with the enzyme gastric alcohol dehydrogenase (ADH), which helps to breakdown the alcohol. Gastric ADH activity is lower in women than in men; one study found that for a given alcohol dose, men's ADH levels were two times higher than women's, and in turn, women's blood alcohol levels were higher than those of men.

Genetic Risk

Family history, adoption, and twin studies all suggest that genetic factors play a role in alcohol use problems and disorders. Some studies have suggested that this role is greater for men than women, although not all do. For example, some twin studies find stronger effects

of genetic factors in the risk of alcohol use disorders in men than women, although others find no gender difference or even a modestly higher heritability for women than for men. A review of twin studies of the genetic influences on alcohol consumption patterns (rather than alcohol use disorders) by Andrew Heath concluded that genetic factors are as important in women as in men for measures of frequency and quantity of consumption, typical weekly consumption, and frequency of excessive drinking, usually defined as having five or more drinks in one drinking bout.

Inconsistent results across studies may be due to methodological differences. Several studies have had relatively small samples of women with alcohol-related problems, due in part to the lower prevalence of these problems in women, and thus have had low statistical power to test hypotheses. Others have relied on records of hospitalization for alcohol use disorders or other social indicators of severe alcohol problems (e.g. drunk-driving arrests) to identify alcohol-related problems in study participants. Such methods detect only the most severe alcohol-related problems, which are more likely among men. Other studies have focused on treatment-seeking individuals, when women are less likely to seek treatment for alcohol-related problems than men.

Summary of Biological Contributors

Although genetic factors could contribute to greater involvement in alcohol for men than women, the evidence is inconsistent. More likely, the greater effects of a given dose of alcohol on women's blood alcohol levels and intoxication compared to men's likely play a role in women drinking less than men.

PSYCHOSOCIAL CONTRIBUTORS TO THE GENDER DIFFERENCE

Cross-cultural differences in the amount of alcohol use in men and women and the size of the gender difference suggest that psychosocial factors may also play a role. The gender difference in alcohol use has been attributed to differences in social norms for alcohol use in men versus women and gender differences in the presence of psychological risk factors for alcohol misuse.

Gender Norms and Roles

Social norms in most cultures permit alcohol use more for men than for women. For men, drinking can serve as a demonstration of their masculinity, their nonconformity, their willingness to take risks, and their superior status. Women who drink, however, are viewed in some cultures as less moral and responsible, more sexually promiscuous, and more sexually vulnerable. Women's drinking is often seen as a threat to their roles as mothers and as the purveyors of "family values." Women are aware of these norms and, to a large extent, enforce them. In a US national survey, 65% of the women said they strongly disapproved of a woman getting drunk, while 58% disapproved of a man getting drunk. In studies in which participants read vignettes about men or women drinking alcohol or cola in the context of a heterosexual date, women drinking alcohol are rated by participants as more sexually available and aggressive than women drinking cola, but no differences in judgments are made about the males in the vignettes. Recent studies using implicit measures instead of self-reports also find that men view women as more sexually available if they are drinking than if they are not.

Ethnographic studies suggest that in cultures in which gender roles are more traditional and pronounced, the gender differences in drinking behavior are greater. Similarly, in US ethnic minority groups that more widely accept traditional gender roles, such as Hispanics and recent Asian immigrants, the gender gap is greater than among European Americans, due largely to high percentages of minority women who completely abstain.

Further, women who endorse traditionally feminine traits (nurturance, emotional expressivity) report less quantity and frequency of alcohol use and adolescent girls and young women who hold more traditional gender role attitudes are less likely to drink at all. On the other hand, males who hold traditional beliefs toward male and female roles are more likely to use alcohol and to show heavy drinking and drinking problems. A study by Rebecca Huselid and Lynne Cooper found that gender-role attributes and ideologies statistically accounted for the relationship between gender and measures of alcohol use. Gender roles completely mediated the gender differences in drinking to intoxication, and partially mediated gender differences in quantity consumed, frequency of heavy drinking (five or more drinks per occasion), and drinking problems. Thus, the theories that differences in social norms and gender roles contribute to the gender differences in drinking behavior have been supported in a variety of studies.

Impulsivity, Sensation-Seeking, Behavioral Undercontrol, and Antisociality

Four personality characteristics that overlap conceptually and cluster together empirically – impulsivity, sensation-seeking, behavioral undercontrol, and antisociality – are consistently related to heavy alcohol

consumption and alcohol-related problems. Males generally score higher on ratings of all these characteristics than females. Thus, the gender difference in alcohol use and problems may be due to gender differences in these personality risk factors.

Further, there is some evidence that this cluster of characteristics is more strongly related to alcohol-related problems in men than women, although the evidence is mixed. For example, some cross-sectional and longitudinal studies have shown that behavioral undercontrol predicts heavy drinking and alcohol-related problems more in men than in women. In contrast, a large twin study found that behavioral undercontrol was equally correlated with lifetime history of alcohol use disorders in men and women, and that behavioral undercontrol accounted for significantly more of the genetic risk for alcohol use disorders in women than in men.

The Dunedin study, which followed youth from early childhood through young adulthood, found that ratings of antisociality when the participants were young adolescents predicted symptoms of alcohol dependence at age 21 for both the male and female participants. Similarly, both males and females who had been diagnosed with conduct disorder between 11 and 18 years of age were more likely than those not diagnosed with a conduct disorder to have symptoms of alcohol dependence at age 21. Antisociality may be more consistently related to alcohol use and problems in males and females than behavioral undercontrol because it results in more severe dysfunction and encompasses a broader syndrome of adolescent problem behavior that includes other drug use, sexual activity, and delinquent and aggressive behavior.

Drinking Motives and Expectancies

People vary in their motives for drinking and expectancies about the effects of alcohol, and these motives and expectancies predict alcohol use and problems. Drinking to cope with distress or depression or to escape from concerns is associated with heavier drinking and more drinking-related problems. Men are more likely than women to report drinking to cope with negative affect or to avoid problems, and drinking to cope or avoid is more strongly related to alcohol use in men than in women. More generally, drinking to cope with stress may be part of a larger avoidant coping style, which is more common in men than women.

Similarly, men are more likely than women to have positive expectancies that alcohol will reduce tension or negative affect, and these expectancies are related to greater drinking. In contrast, laboratory studies show that women are more likely than men to expect alcohol to interfere with their ability to cope with difficult situations, and to avoid alcohol when they must deal with stressful situations.

Summary of Psychosocial Contributors

Drinking, particularly heavy drinking, is more socially accepted for men than for women, and more a part of the male gender role. These social norms and roles appear to account substantially for the gender difference in drinking and related problems. In addition, men tend to carry psychological risk factors for problematic drinking more than women, including impulsivity, sensation-seeking, behavioral undercontrol, antisociality, and drinking to cope with stress.

CONSEQUENCES OF MEN'S AND WOMEN'S DRINKING

The physical health and social consequences of drinking may differ for men and women, contributing to the gender differences in drinking behavior.

Physical Health Consequences

Among people who develop alcohol use disorders, the progression from first use to these problems appears to be faster for women than for men, a phenomenon known as *telescoping*. For example, retrospective studies of alcoholics suggest that women progress from first getting drunk regularly (e.g. weekly) to a diagnosis of alcoholism more quickly than men do.

Perhaps because of the increased bioavailability of alcohol in women compared to men, women may suffer more physical diseases as a result of alcohol use compared to men. A meta-analysis of 38 studies found that women are more likely than men to report health problems when they have a history of heavy drinking, alcohol abuse, or alcohol dependence. Similarly, another study found that women's risk of death increased by 160% if they were heavy drinkers (defined as six or more drinks per day) compared to if they were light drinkers (more than one drink per month but less than one drink per day), whereas heavy drinking increased men's mortality risk by only 40% over light drinking.

Moderate alcohol use is associated with lowered risk of myocardial infarction or cardiac death in men and in postmenopausal women, but excessive alcohol consumption is associated with a significantly increased risk of cardiac diseases, including cardiomyopathy, dysrhythmia, and hypertension. Women may show cardiac harm at lower levels of drinking than men. Analyses of a survey of 43 763 men and women showed that among women, even moderate drinking (defined as

more than two drinks per day during their heaviest drinking period in the last year) was associated with a significant increase in risk for heart disease. Among men, heavier drinking (more than five drinks per day in their heaviest drinking period in the last year) was associated with an increased risk for heart disease.

Women who are dependent on alcohol have 50–100% higher rates of premature death than do men who are dependent on alcohol. Women dependent on alcohol show comparable rates of diseases associated with alcohol dependency as do men dependent on alcohol, such as ulcers, gastrointestinal hemorrhage, fatty liver, hypertension, obesity, anemia, and malnutrition, even though the women tend to have shorter durations of drinking.

Women with diagnosed alcohol use disorders are more likely than women without alcohol use disorders to experience sexual dysfunctions, cessation of menstruation, irregular menstrual cycles, and early menopause. Alcohol use is a risk factor for breast cancer in women.

Heavy drinking by pregnant women is associated with fetal alcohol syndrome, which is characterized by growth deficiency, altered structural development, and central nervous system dysfunction due to the exposure of a fetus to alcohol in the womb. Even low to moderate levels of drinking during pregnancy have been associated with subtle abnormalities in reproductive outcomes such as birth weight, gestational age, rate of miscarriage or stillbirth, congenital abnormalities, and social and cognitive development. Many women stop drinking when they are pregnant and do not resume heavy drinking after they give birth, in part because of the demands of parenting small children.

In men, alcohol reduces testicular function, but whether it affects reproductive outcomes or infant development is unclear. Both animal and human studies have produced inconsistent results on whether paternal drinking around the time of conception has effects on fertility or fetal outcomes.

The greater physical health effects of drinking on women compared to men may lead more women to stop drinking or to drink more lightly than men, particularly when pregnant and in older age.

Violence and Sexual Assault

The risk of sexual assault and violence to women is substantially increased when they are under the influence of alcohol. A longitudinal study found heavy alcohol use predicted later physical assault victimization for women but not men. Women who abuse alcohol (or other drugs) are at particularly high risk for sexual or physical assault. More generally, being under the influence of alcohol appears to increase a woman's risk of assault by her male partner. One study of over 1000 women who had been sexually assaulted found that over half the assaults involved alcohol; in almost all cases, the man had been drinking, and in the majority of cases the man and woman had both been drinking. Women may become more physically abusive when they are drinking, but women's assaults most often come in the context of drinking by both the woman and the man, leading to mutual aggression toward each other.

Risky Sexual Behavior

One of the fears concerning women drinking is that it will make them sexually promiscuous. General surveys of the population suggest that adults who use more alcohol are more likely to report more liberal sexual attitudes and to report participating in risky sexual behavior, such as unprotected sex, sex with multiple partners, or sex with partners at high risk for sexually transmitted diseases. Similarly, surveys of adolescents and young adults have found that those who use more alcohol report more risky sexual behavior. These links may not be due to the alcohol itself, however, but to a third variable of sensation-seeking, behavioral undercontrol, or impulsivity.

Indeed, some studies find that women become more conservative in their sexual behavior when they are drinking. Studies of adolescents' first dates with an opposite sex partner find that girls are less likely to agree to intercourse if they have drank alcohol on the date, suggesting that alcohol inhibits girls' willingness to engage in sex with males they do not know well. Boys, however, reported that they were more likely to engage in intercourse or heavy petting on a first date if they had been drinking.

In experimental studies, men and women are told they will receive an alcohol beverage or a placebo; this is crossed with the actual administration of an alcohol beverage or placebo, creating four groups. Men who believe they have consumed alcohol become more sexually aroused and forward in their interactions with women, regardless of whether they have actually consumed alcohol or not. Women, however, do not show consistent effects of expectations, and sometimes show suppression of sexual arousal and interactions with male partners.

Summary of Consequences

Both men and women may suffer physical health problems, violence and sexual assault, and engage in riskier sexual behavior when drinking, particularly when drinking heavily. There is some evidence that women experience more of these negative consequences, and at lower levels of drinking than men.

GENDER AND TREATMENT FOR ALCOHOL USE DISORDERS

The treatment of alcohol use disorders is challenging. Meta-analyses and reviews of existing treatments suggest that, as a whole, they help only about 17–35% of substance-dependent individuals abstain for up to 1 year.

Women with alcohol use disorders are less likely than men with alcohol use disorders to seek treatment. This may be due in part to the greater stigma against excessive alcohol use in women. Women also are more likely than men to have child care responsibilities, and transportation and financial problems that impede treatment-seeking.

Some studies find that women may respond better to women-only treatments than to mixed-sex treatments. In addition, treatments that address common comorbidities with alcohol use disorders in women, especially depression and posttraumatic stress disorders related to sexual assault, are associated with fewer relapses and higher treatment satisfaction.

CONCLUSIONS

Men appear to be more likely to carry a number of risk factors for alcohol use and alcohol-related problems. Some studies suggest that genetics may play a stronger role in alcohol use problems in men than women, although the literature is somewhat inconsistent. Women show higher blood alcohol levels and become more intoxicated at lower levels of alcohol intake, which may protect them from heavy drinking and building tolerance.

Drinking is more socially acceptable for men than women, and men are judged less harshly for heavy drinking and intoxication than are women. Men who endorse traditional gender roles are more likely to drink while women who endorse traditional gender-stereotyped traits are less likely to drink.

Men are more likely than women to have a number of individual difference characteristics associated with alcohol use and related problems. These include impulsivity, sensation-seeking behavioral undercontrol, and antisociality. Men are also more likely than women to drink to cope with distress and to have positive expectancies for the tension-reducing effects of alcohol, which contribute to more drinking.

The consequences of alcohol consumption are different for men and women. Women progress to alcohol use disorders given heavy drinking faster than men, suffer alcohol-related physical illnesses at lower levels of exposure to alcohol than men, and heavy alcohol use is associated with several reproductive problems in women. Women may also be more likely to suffer physical harm and sexual assault when they are using alcohol.

Both the absence of risk factors and the greater negative consequences of alcohol use may play roles in women drinking less than men and having fewer alcohol-related problems than men. Women may notice they are becoming intoxicated at a much earlier stage of intoxication than men and may be more likely to find these effects aversive or frightening, leading women to inhibit their alcohol consumption. This, in turn, may protect women from developing tolerance to high doses of alcohol, and alcohol-related social and occupational problems.

If alcohol consumption reduces women's reproductive capacity, and leads them to produce offspring that are less likely to survive and to produce their own offspring, then there would be selection pressures against alcohol consumption in women. Women who were more sensitive and averse to the early negative effects of alcohol on their functioning might be less likely to develop alcohol-related physical diseases and more likely to produce healthy offspring than women who were less sensitive and averse to the effects of alcohol.

Women with alcohol misuse problems are less likely to seek treatment than men with alcohol misuse problems. Decreasing barriers to treatment-seeking (e.g. providing child care) and women-specific treatments may increase women's help-seeking.

SEE ALSO

Alcohol Use Disorders, Cognitive Factors in Addictive Processes, Emotions and Addictive Processes, Contextual Factors in Addiction, Epidemiology of Addiction, Families and Addiction, Impulsivity, Disinhibition, and Risk Taking in Addiction, Interpersonal Factors and Addictive Disorders, Maturing Out, Peer Influences on Addiction, Personality and Addiction Processes, Substance Use and Mood Disorders, Alcohol and Sexual Violence

List of Abbreviations

ADH	alcohol dehydrogenase
GENACIS	Gender, Alcohol, and Culture: An International Study
NCS	National Cormorbidity Survey
NLAES	National Longitudinal Alcoholic Epidemiologic Survey

Glossary

Alcohol abuse DSM-IV diagnosis defined as persistent drinking behavior in the face of repeated social, interpersonal, and occupational problems due to excessive alcohol consumption.

Alcohol dependence DSM-IV diagnosis that includes the psychosocial problems involved in alcohol abuse, but can also involve physiological dependence on alcohol, such as tolerance and withdrawal symptoms. Also referred to as *alcoholism*.

Alcohol misuse maladaptive pattern of drinking that may qualify for diagnoses of alcohol abuse or dependence.

GENACIS project Gender, Alcohol, and Culture: An International Study; a multinational study of alcohol use and correlates.

Heavy drinking or binge drinking drinking a large amount of alcohol in a short period of time, often defined as five or more drinks in 2 h for men and four or more drinks in 2 h for women.

National comorbidity survey Large-scale survey of psychiatric problems in the United States.

Further Reading

Nolen-Hoeksema, S., 2004. Gender differences in risk factors and consequences for alcohol use and problems. Clinical Psychology Review 24, 98–101.

Stewart, S., Gavric, D., Collins, P., 2009. Women, girls, and alcohol. In: Brady, K.T., Back, S.E., Greenfield, S.F. (Eds.), Women and Addiction: A Comprehensive Handbook. Guilford Press, New York, pp. 341–359.

Wilsnack, R.W., Kristjanson, A.F., Wilsnack, S.C., Crosby, R.D., 2006. Are U.S. women drinking less (or more)? Historical and aging trends, 1981–2001. Journal of Studies on Alcohol 67, 341–348.

Wilsnack, R.W., Wilsnack, S.C., Kristjanson, A.F., Vogeltanz-Holm, N.D., Gmel, G., 2009. Gender and alcohol consumption: patterns from the multinational GENACIS project. Addiction 104, 1487–1500.

Relevant Websites

http://www.niaaa.nih.gov – National Institute for Alcoholism and Alcohol Abuse.

http://orwh.od.nih.gov – Office of Research on Women's Health of the National Institute of Health.

http://women.webmd.com – WebMD Home Page for Women's Health.

CHAPTER

16

Minority Groups and Addictions

Arthur W. Blume

Washington State University, Vancouver, WA, USA

OUTLINE

Minority populations often have different health and mental health needs and outcomes than the majority. With regard to substance use, minority groups within societies have been found to have different substance use presentations and responses to treatment than the majority group in the same society. Group differences noted by researchers are generally attributable to cultural differences between minorities and the majority and to differential treatment in society.

DEFINITION OF A MINORITY GROUP MEMBER

The term minority group seems to imply a population subgroup numerically smaller in size than a majority group in a culture, but such a definition does not account for being politically disempowered. In reality, minority groups can be the largest group in a society numerically, but not wield dominant political power. Because of the unusual circumstances that occurred during the empire building colonization by European nations in many areas of the world, European colonialists have exercised significant political power in many societies in spite of being a numeric minority.

The majority group regardless of its numbers has the majority of political power and determines the cultural norms, practices, and worldview of a given society, whereas minority culture is a subgroup with a history of being disempowered in society that usually have significantly different cultural norms, practices, and worldview than the majority culture. A minority group member would be a person who self-identifies with (or is identified by others as belonging to) a minority group. Minorities, because of their different assumptions about the world than the majority, have sometimes been oppressed by the majority for their views.

The focus here is on minority groups defined by ethnicity/race, but it is worth noting that there are other minority groups in society defined by gender, sexual

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00016-4

orientation, and disability that also have different substance use presentations and treatment needs than the majority group within particular societies.

HISTORICAL CIRCUMSTANCES

The historical circumstances of minority populations within a particular society are often different than the majority within the same society. Because the historical contexts for minority groups differ from majority group members, they often have unique health and mental health issues not shared by the majority. Minority groups have experienced disempowerment and oppression at the hand of the majority in historical power struggles within many societies. The causes for majority–minority group conflict are many, but include historical feuds between groups, prejudice and racism, and economic motives.

Colonialism

Colonial expansion by Europeans into Africa, Asia, Australia, the Caribbean and Pacific Islands, and North and South America profoundly shaped minority experiences in these areas of the world. Europeans were eager to move to new territories opened to them by exploration to seek freedom and economic opportunities. During the mass migration of people, European settlers found the areas occupied by indigenous people. First encounters varied from peaceful to violent depending on circumstances, but soon armed resistance by some indigenous people to what they perceived to be an invasion of their historical territories became more prevalent.

The interaction of Europeans to indigenous people also included exposure to pathogens not previously experienced by either population. An exchange of infections occurred in both indigenous and Europeans, but the clear losers in the exchange of pathogens were the indigenous people who did not have immunity to serious and rapidly spreading diseases for which Europeans had developed some resistance. The most famous exchange of pathogens occurred in the Western hemisphere in what was called the Columbian Exchange when Indians and Natives were plagued by diseases such as measles, small pox, and tuberculosis for which they had no natural immunity. Some demographers and public health scientists have estimated that 80–90% of exposed indigenous populations died during these epidemics, leaving behind vast areas of abandoned or under populated areas that were often noted in historical accounts by European settlers. The decimation of many communities and civilizations left indigenous societies vulnerable to conquest. Illnesses together with the superior weaponry of European soldiers (in spite of being outnumbered) contributed to lopsided victories for the colonists.

With regard to substance abuse, European colonization brought the introduction of distilled spirits into indigenous cultures. Although many indigenous cultures had fermented grain and berries for special ritualistic use in religious practices and medicines, many groups had not developed distillation methods. For many communities, the introduction of potent distilled alcohol caused significant behavioral and health problems. Europeans also used alcohol as an everyday drink and for social occasions rather than religious or spiritual practices. Daily use was unfamiliar to many indigenous groups prior to exposure to European colonists. Within a generation or two, substance use patterns among indigenous communities shifted from special and occasional to regular and excessive, and to the use of potent distilled liquor rather than fermented beers and wine. The results were devastating to many communities. Some European colonists used the knowledge that indigenous communities were having difficulties with substance use to their political advantage. For example, distilled spirits were used by various colonial powers to influence treaty negotiations with indigenous people, impairing the judgment of tribal representatives. The result was loss of access to traditional territory and forced removal of various indigenous groups, as well as tribal infighting that further weakened indigenous communities as a result of substance use and community anger at the concessions made by various leaders.

Derivative Eurocentric societies were spawned from the legacy of European colonialism in many areas of the world. The indigenous populations within those Eurocentric cultures were subsumed into the role of minorities, even though in some cases they represented the numeric majority. Because indigenous cultures were generally conquered people, they almost without exception had little political power in colonial societies for years, and in some cases, the political imbalances continue today. The consequences of these imbalances have contributed to a number of risk factors for substance abuse among minority groups in these societies, including the experience of oppression and discrimination, trauma, poverty, and lack of access to quality health and mental health care. These risk factors for minority substance abuse will be discussed subsequently.

Slavery and Indentured Servitude

For many minority groups, being enslaved or used as indentured servants by the majority culture has been an important part of group identity in the society in which

they live. Slavery has been in existence for a significant portion of human history. Although slavery is often a function of a majority group enslaving members of a minority group, some minority groups have a history of slave ownership themselves. In the new world, European soldiers enslaved indigenous people to help them build their settlements. However, the diseases of the Columbian exchange so severely reduced the native population that in the estimation of European colonists there was not enough indigenous labor available to fill the labor needs. Hence, Europeans began the slave trade and imported African slaves into the new world, a practice that lasted for over 300 years in many societies.

Indentured servitude, on the other hand, is a practice in which a person agrees to labor for another person for a specified length of time in exchange for immigration into a society, transportation, food, shelter, use of land, or other possible material benefit. The idea was that a person who otherwise could not afford something would trade hard work for the item being sought. The reality was that indentured servitude was a way to enslave people for significant periods of time with the promise of an eventual material reward. Those who employed indentured servants often found ways to compound the debt by underhanded means, thereby extending the time of service exponentially. Many minorities could not afford immigration and would agree to indentured servitude as a means of immigrating into a society. Indentured servitude also has been used as a way to continue to deny economic and political power to minorities, especially those who were emancipated from slavery (e.g. sharecropping).

The consequences of slavery and indentured servitude have contributed to increased risk factors associated with substance abuse, including poverty, health and mental health problems, and lower educational attainment among minority groups over many generations. In societies where slavery existed for centuries, the intergenerational consequences of oppression have been difficult to remediate.

Assimilation Efforts

Majority groups sometime pressure minorities to conform to majority group expectations and norms. The European derivative societies that resulted from colonization engaged in this practice to a varying degree, often as a function of the perceived threat of a minority group to the majority culture and of the level of ethnic superiority practiced by the majority culture. The expectation that minorities will discard their minority culture in favor of embracing majority culture beliefs and practices is referred to as assimilation. The goal of assimilation is to extinguish cultural differences to achieve a homogeneous society. Minority groups are expected to discard their native language, traditional cultural practices, traditional spiritual practices, and any other beliefs or practices that are discordant with majority beliefs and practices. Some European derivative societies, including the United States, had written policies directing organized assimilation efforts during colonial periods.

In the United States, policies toward assimilation of American Indians and Alaska Natives were implemented in conjunction with military conquest campaigns. American Indians and Alaska Natives were promised rewards in treaties for giving up traditional practices. Many children were removed from families and sent to boarding schools that were in some cases hundreds of miles from their homes. In these boarding schools, speaking the native language or engaging in traditional practices was punished and education was designed to indoctrinate children into American society. In another government effort, native people were lured away from reservations into urban centers with promises of jobs in efforts to assimilate them into the larger society. The result of assimilation was a general disintegration of traditional culture and family structures in American Indian and Native Alaska communities. It has no coincidence that substance abuse became a problem for many of these communities during this time. Many American Indians and Alaska Natives refer to conquest and assimilation efforts as a "soul wound" which they directly attribute as the cause for a number of health and mental health problems including substance abuse.

Isolation

Some minority communities have experienced isolation as a consequence of historical circumstances. In some instances, isolation has been imposed on minority groups by the majority culture. Examples of imposed isolation include relocating indigenous people into special reserves that are often located in sparsely inhabited areas, the legal segregation of communities defined by skin color or ethnicity, and the internment of certain groups in times of war or threat of war. In other instances, isolation has been self-imposed by the minority communities, often as an effort to protect traditional practices. Self-isolation often results in enclaves of minority communities operating independently of the majority culture and interacting with the majority culture when it is absolutely necessary. Cultural isolation has contributed to a number of risk factors related to substance use such as poverty, poor access and quality of health and mental health services, poor education outcomes, and societal neglect.

CULTURAL FACTORS ASSOCIATED WITH SUBSTANCE USE PATTERNS AND TREATMENT OUTCOMES

Differences in Worldviews

One result of European colonization has been the development of pluralistic societies that have brought together a wide range of cultural worldviews. Basic tenets of European culture such as individualism, personal autonomy, and the linear flow of time are not shared by the minority cultures they conquered. Assimilation efforts, although comprehensive in many societies, did not succeed in eliminating all divergent worldviews. In addition, many societies have experienced significant post-colonial immigration with improvements in transportation. One source of significant migration in Europe, for example, has been by immigrants from former colonies, rapidly diversifying many European nations that historically had been rather homogeneous.

Differences in cultural worldviews have been identified as one key variable to consider when studying and treating addictive behaviors among minorities using substances. Because worldviews of minority groups often differ radically from majority culture, one cannot assume that majority culture approaches to investigating and treating addictive behaviors will work effectively among minority populations, and in fact, there is able evidence to the contrary.

Cultural differences between majority and minority cultures vary by society. It is hard to make generalizations about cultural differences for the many societies where European colonialism occurred, but there are often major differences across cultures. For example, in many European derivative societies, the view of time is linear and view of space is that it includes the observable and measureable world. In some minority cultures within those Eurocentric societies, time is viewed as cyclical and repeating and space is viewed as visible and invisible (the natural and supernatural coexist).

Some minority cultures hold more collectivistic views, although this is not true for all. The welfare of the community is more important than the welfare of the individual, and in fact, the two may be viewed as inseparable. Many minority cultures view family roles and structures as more important than individual identity. Some also may believe in fate and destiny over free will and free choice. Autonomy is not something that is either desired or sought in some minority cultures. In some minority cultures suffering is at the center of understanding existence whereas in many European derivative cultures suffering is something to be avoided or overcome.

Understanding cultural beliefs is essential for conducting competent addictive behaviors research in minority communities. Since much of addictive behavior research is formulated, implemented, and analyzed by majority group scientists, they often lack awareness and understanding of the cultural differences in minority groups that may impact their research. Important culturally bound constructs within a minority culture that are not similarly present in majority culture may be inadvertently ignored in research, and scientists, assuming construct equivalence, may mistakenly assume that important addictive behavior variables of interest in general population studies exist similarly in the minority culture. Another difficulty is that scientific methods common deployed with success in majority culture may conflict with minority culture beliefs and practices, contributing to distrust of addiction scientists among minorities and poorly conducted research.

Understanding minority cultural beliefs is also essential to providing effective treatment services to minority clients. It would be extremely difficult to develop the therapeutic alliance with a minority client without understanding the cultural issues that influence the thinking and behavior of the client. It would be difficult to aid a client without understanding the cultural roles and expectations for that client. It also would be difficult to develop an effective treatment plan without understanding the everyday challenges that a minority client may face in her or his society.

Acculturation and Enculturation

Acculturation and enculturation levels have been found to be important predictors of the therapeutic alliance and for treatment outcomes for minorities engaging in the addictive behaviors, and are critical variables to account for when developing effective treatment plans. Acculturation level is the degree to which a person from a minority group functions and interacts competently within the majority culture. High levels of acculturation are associated with high levels of competence in understanding and using majority culture beliefs and practices and engaging successfully in majority culture functions. Highly acculturated minorities often are proficient at use of majority language, have received advanced education in majority culture schools, employed by majority culture entities, and are able to function competently on a daily fashion in the majority culture. Low levels of acculturation, on the other hand, would be associated with little engagement in majority culture activities, little understanding of majority culture beliefs or practices, and sometimes little use of majority culture language. People low in acculturation often have not received education from a majority culture institution, therefore are often at a competitive

disadvantage for employment opportunities in majority culture.

Enculturation level is the degree to which a person from a minority group functions and interacts competently within her/his minority culture. High levels of enculturation are associated with high levels of competence in understanding and using minority culture beliefs and practices and engaging successfully in minority culture functions. Highly enculturated minorities are proficient at use of her/his ethnic language and are able to function competently on a daily fashion in the minority culture. Low enculturation would be associated with little engagement in minority culture activities, little understanding of minority culture beliefs or practices, and little use of minority culture language.

Researchers have found that having high levels of both acculturation and enculturation is a protective factor for minorities from addictive behaviors. Having high levels of both acculturation and enculturation has been referred to as bicultural competence for personal expertise in majority and one minority culture, or multicultural competence for personal expertise in majority and more than one minority culture. Many culturally relevant addictive behaviors prevention and treatment programs used with minority clients include skill training components to enhance bi- and multicultural competence in participants to improve outcomes.

Ethnic Identity

Ethnic identity is self-identified ethnic group membership. Ethnic identity evolves from highly personal beliefs that reflect personal experience and social interactions. Ethnic identity is highly variable even among people who look and think alike. It is important to note that members of minority groups do not always self-identity with being a minority on a daily basis, and in some cases due to assimilation or low enculturation, may actively resist being identified as a minority group member. The difficulty for such a person is when they are identified as a minority group member by others, especially if such identification happens in ways that are uncomfortable or hurtful. Such an event can create cognitive dissonance between personal beliefs about ethnic identity and societal beliefs about group assignment.

Ethnic identity has been used to identify appropriate addictive behavior treatment and prevention strategies. Some research suggests that an ethnic identity (like to like) match for choice of treatment and prevention methods may improve outcomes. Certainly, it is important for clinicians to recognize and respect the wide variations in ethnic identity that minority clients bring to therapy in order to enhance the therapeutic alliance. Research also highlights the importance in avoiding assumptions about ethnic identity.

Discrimination, Prejudice, and Racism

Research has found that minority groups face discrimination, prejudice, and racism routinely with detrimental effects. Discrimination involves treating a person differentially and in a way detrimental to the person because of bias toward that person's group membership. Prejudice involves biased attitudes about groups of people that are insulting or injurious to a minority group member that are often expressed or acted upon automatically and without awareness. The attitudes are often the result of stereotypes about a particular group that is pervasively held in majority culture. Stereotypes are broad swath generalizations about group members (in this instance, minority group members) unsupported by empirical evidence. Stereotypes are perceived as insulting by minority group members even though they are sometimes perceived as flattering by the majority culture (even though untrue). Stereotypes are the net result from divisive us versus them thought processes.

Racism is rooted in belief in biologically based group differences (often defined by skin color and phenotypes) that causes certain groups to be biologically and intellectually superior to another. Racism also includes stereotyping, but the assumptions are that stereotypes reflect real biological differences between groups. Racism generally contributes to beliefs in the superiority of majority culture because of innate racial hierarchies. Racism also often contributes to the belief that it is important to maintain distances between racial groups, often to maintain racial "purity," and generally results in hostile thoughts or actions against minority group members. Racism can be expressed overtly (highly visible and aggressive) or covertly (veiled and done passive-aggressively).

Minority group members frequently report the experience of discrimination, prejudice, or racism. Public health researchers have linked these regular experiences to increased stress, poor overall health outcomes, and increased substance abuse. As an example, researchers have found that minority group members often experience subtle stereotyped insults or put downs on a daily basis, referred to as microaggressions. Microaggressions can be verbal (e.g. an insulting comment), written (e.g. an offensive statement), or visual (an offensive image). Perpetrators of microaggressions often lack awareness that their comments or actions are disturbing and hurtful. Microaggressions are likely the result of implicit racial and ethnic bias that permeates a society (see following section). Psychologist Derald Wing Sue has written a great deal about the different kinds of

microaggressions found in American culture and their consequences for people of color.

SPECIAL CONSIDERATIONS FOR RESEARCH AND CLINICAL PRACTICE

Cultural researchers have found consistent evidence for implicit bias against minorities in society. From hiring practices to leadership styles, experimental research has demonstrated bias in interpreting behavior and making hiring decisions that favor majority culture members. In addition, social psychologists have found substantial evidence documenting implicit bias that affects perceptions of others, attitudes, and decision-making in people without their awareness. Even minority group members have been found to often have implicit negative bias against their own groups. The balance of evidence suggests that bias against certain minority groups is present and pervasive in all members of society regardless of group membership, and because of that bias we cannot trust our objectivity.

Assessment

Scientists are creatures of culture as are all humans. Although researchers strive to be completely objective, it is simply impossible to divorce one's self from one's socialization. In spite of empirical training, culture impacts the development of theoretical models to study research questions, decisions about how to investigate research questions, and interpretation of results.

Most of the gold standard measures used in the research of addictive behaviors have been developed in the context of implicit majority culture bias. They are constructed from models that typically reflect "Western" ways of thought that include basic assumptions about the world that may not be accepted by certain minority groups. Many behaviorally oriented measures use majority culture language or were first developed in majority culture language and translated. Sometimes the language also reflects worldview biases. Many gold standard measures of addictive behaviors have been developed with majority culture assumptions that likely do not apply to certain minority cultural worldviews. Construct equivalence across cultures has been assumed in the development and use of these measures, but the reality is that vast cultural differences make the assumption of construct equivalence flawed. In addition, many of the quantitative measures have been inadequately normed for many minority cultures and therefore contribute to questionable interpretation of results.

Diagnosis

Since assessment tools are generally biased, the ability to make an accurate diagnosis is impacted. Indeed, public health researchers have noted that mental health diagnostic assessment tend to overpathologize minority respondents through artificially inflated symptom scores due to measurement error or by misinterpreted results. In addition, minority group members have been found to describe symptoms differently than typically observed in majority, causing health care professionals to miss important information about symptoms and contributing to missed diagnoses and prescribing errors. As an example, there is no exact translation of the English word binge into Spanish, so clients whose first language is Spanish will not naturally use the term to describe substance use behavior. Treatment professionals should be aware that some minority members may describe their addictive behaviors and subsequent consequences differently than majority group members.

Minority cultures also may have culturally bound conditions that treatment professionals may not know. Sometimes a culturally bound condition is linked to the addictive behavior being assessed and treated. It is important for treatment professionals working with minority clients to understand culturally bound conditions and their association to addictive behaviors in their clients. Some cultures view addictive behaviors as symptoms of a more pervasive culturally bound condition. As an example, some American Indians and Alaska Native groups will refer to problems with substance abuse (and other health and mental health concerns) as a result of a soul wound that has resulted from collective trauma experienced across the generations as a result of colonial conquest and assimilation efforts. Under those circumstances, a minority client will believe that treating the culturally bound condition will impact not only the addiction but also perhaps multiple health and mental health problems simultaneously. In other words, culturally bound conditions would be considered systemic and the primary condition to be treated, whereas the addictive behavior would be considered a symptom of the underlying condition.

Stereotype Threat

Research by Claude Steele and others have found that minority group members may experience stress that can negatively impact outcomes when they are engaged in a performance task involving a stereotyped behavior for their group. This phenomenon is referred to as stereotype threat. For example, if a stereotype exists in majority culture that suggests a member of a minority group will perform poorly on a test or task, the expectation by

society of poor performance places undue stress and anxiety on the minority group member that can actually contribute to poor performance. When assessing and treating the minority group members for addictive behaviors, clinicians should be aware of the possibility of stereotype threat that might contribute to poor performance and outcomes in treatment. A savvy professional will want to determine if stereotype threat is impeding progress.

Ethics and Research among Minorities

Researchers have made a number of missteps when investigating addictive behaviors among minority communities. Perhaps the most notorious of studies is referred to as the Barrow Alcohol Study, which has been written about extensively. The study was conducted in Barrow, Alaska, in the late 1970s. The study was requested by public health officials in the Barrow area that had concerns about alcohol abuse among the Inupiat people of Alaska. The public health group sought help from investigators in Pennsylvania with no experience in working with Native Alaska populations. The investigators developed and implemented the study methods without any consultation with Inupiat community members. Data collection focused solely on drinking behavior of some members of the Inupiat community without providing appropriate context for the drinking behavior by including non-native members of the Barrow community as a comparison group in the study. The measures used were not normed or tested in the Inupiat community prior to the study, and included measures not representative of the worldviews of Inupiat people. The results of the study were released to the press without community input or knowledge, and resulted in an embarrassing and damaging story that appeared in national media venues including the *New York Times*. The press release was meant to coerce the Inupiat community into changing drinking behavior, but instead resulted in Inupiat community outrage and increased distrust of majority culture researchers by American Indians and Alaska Natives in Barrow and beyond.

Researchers of addictive behaviors investigate issues of a highly sensitive nature given the stigma attached to addictive behaviors in many societies. Because of the history of societal disempowerment and oppression, and a track record of poorly conducted research that have shamed communities, there is a burden for proof on addictive behavior researchers to demonstrate extra care in conducting research with vulnerable minority communities. In addition, ethical violations have made conducting minority research a difficult enterprise because distrust of scientists is deservedly high.

Ethics and Clinical Practice among Minorities

Multicultural competence is an important skill to develop to effectively treat the addictive behaviors of minority group members. Some clinically oriented professional organizations have instilled training requirements for multicultural cultural competence into graduate training programs (e.g. the American Psychological Association) because of the research findings suggesting its importance for improved clinical outcomes. Treatment programs for addictive behaviors that serve minority communities may expect their therapists to receive additional training in multicultural methods and often opt to include culturally relevant approaches to treatment in the programs. In some instances, progressively minded service providers have established collaborative relationships with traditional healers and elders form the minority community being served. The result is holistic care that includes the best practices from minority and majority culture.

An appropriate match of client to therapist by common ethnicity has long been considered an important factor for a positive treatment outcome. The research evidence is mixed concerning the efficacy of ethnic identity match for improved outcomes in psychotherapy. What appears to be more important is an appropriate match of therapy and therapist to the acculturation level of the client to improve the therapeutic alliance. For a client low in acculturation, it might be more important to have an appropriate ethnic match than for a client high in acculturation level. Many culturally relevant therapies for addictive behaviors employ the use of acculturation match to potentially improve treatment outcomes.

Empirically Based Substance Abuse Practices for Minorities

Addictive behavior research has used complex experimental methods including randomized controlled clinical trials to validate the efficacy of particular interventions within the general population. However, it is difficult to know if a particular intervention will work similarly with minority groups as it did in general population studies. Many intervention efficacy studies have not had sufficient sampling of some minority groups to be able to determine the efficacy for those group members. Minority group members frequently respond differently to interventions than majority group members. Having insufficient data concerning minority group responses in efficacy trials lead the American Psychological Associations' Task Force on Empirically Validated Therapies to conclude that, "…the efficacy of empirically validated therapies has not been established

with ethnic minority populations." Although a number of highly regarded and empirically supported interventions for addictive behaviors have been used with various minority group participants, there has generally been insufficient research in the area to draw any conclusions about whether established empirically based practices will work effectively and consistently among minority communities.

In the past, researchers have used a utilitarian approach concerning the use of empirically supported interventions, reasoning that what is good for the majority would likely be good for minority groups as well. Initial efforts to transport empirically supported therapies into some minority communities did not attempt to account for cultural differences of the community being served. In discovering that minority participants often did not respond in the same way to the interventions as would be expected in the general population, methods were revised and the process became to modify empirically supported interventions to be more culturally relevant by adding new components consistent with the beliefs and practices of the community. These efforts have met with limited success, and even among those that showed promise, it was unclear which components may have been most effective, the original intervention or the culturally relevant add-ons. In reality, in order to tailor interventions for effective use within minority communities, it would be important to include community stakeholders during initial development of an intervention, rather than in an advising role to assist scientists in making an existing intervention more culturally relevant to the community being served. Today, many researchers working with minority communities to develop effective interventions on addictive behaviors understand it is important to seek early community participation to make certain the intervention is developed with community in mind from the beginning rather than borrowed and revised as an afterthought.

Model for Community-Based Participation

Researchers have to overcome challenges between bridging the culture of science to the culture of the minority community being studied. Since minority community worldviews, beliefs, and practices can differ sharply from the assumptions and expectations of researchers, there is a high potential for mistakes that damage collaborative relationships.

Because of frequent missteps that scientists have made in working with minority groups, as well as serious concerns about the lack of information about what will effectively intervene on addictive behaviors with many minority communities, a new model of scientific inquiry referred to as the Community Based Participatory Research (CBPR) Model has been developed and used with great success. The CBPR model proposes a collaborative relationship between the community being investigated or served and addiction professionals. In the CBPR model, researchers serve as expert advisors to the community on what is known to be best scientific practices as particular studies or services are being developed, implemented, and analyzed. However, community stakeholders will define who will participate, what will be investigated and how the investigation will proceed, and how the results will be disseminated. Addiction scientists advise community stakeholders using professional expertise and scientific judgment in their roles of consultant-scholars, but the ultimate decisions about methods lie with the one's being impacted most directly by the research. High on the list for reasons to use such an approach include conducting research with enhanced cultural sensitivity and increased oversight of methods that will protect the community from harm and embarrassment. Using the CBPR to guide research of addictive behaviors in minority communities has increased trust of minority communities in researchers, provided tangible and more immediate benefits for minority groups, and improved scientific methods for understanding addictive behaviors among minority group members.

HEALTH DISPARITIES AND ADDICTIVE BEHAVIORS

Researchers have found consistent evidence for addictive behavior health disparities among some minority groups. Patterns for health disparities vary across societies. Minority group disparities have been found both in the prevalence of certain types of addictive behaviors and in the accessibility of treatment services. Some but not all minority groups have been found to have significantly higher rates of addictive behaviors than those reported in general population. In addition, some minority groups have significantly lower utilization rates of addiction treatment services than would be predicted by the prevalence rates of the group, and some groups have significantly less utilization of treatment services than observed in the general population.

The contributors to health disparities are many. Greater rates of addictive behaviors are often linked to conditions to which many minority groups are exposed, such as oppression and poverty. Prevention efforts have not been historically targeted within some minority communities with the same intensity as in the general population. There is evidence that specific minority groups have been targeted for intense and culturally specific alcohol and tobacco advertising

and marketing campaigns. In some minority communities, the density of liquor markets, bars, and convenience stores selling legal psychoactive substances have been found to be significantly higher than surrounding majority culture neighborhoods. Financial burdens linked to poverty or unemployment often reduce accessibility to treatment, as does lack of health care insurance that covers addiction treatment services. Access is also restricted in rural locations, by number of local treatment providers, and by accessibility of transportation to treatment facilities. Unfortunately, it is often true that treatment services in remote locations or near impoverished neighborhoods, or provided to people with limited means or the uninsured, are also not always of the quality provided in more affluent and urban areas.

CONCLUSION

Minority clients have unique needs that require approaches that respect cultuniral beliefs and values. Historical circumstances have frequently created disadvantages for minorities and many of these disadvantages remain problematic today. These difficult circumstances have contributed to disparities in the prevalence of and care for addictive behaviors. In the past, addictive behavior researchers made missteps that harmed minorities and produced inadequate data. Addiction professionals have improved their research and clinical practice by enlisting community stakeholders as collaborators. Use of the CBPR model has created opportunities not only to serve minorities more effectively, but also to develop creative and culturally relevant new interventions for addictive behaviors.

SEE ALSO

International Perspectives on Addiction, The Biopsychosocial Model of Addiction, Historical Understandings of Addiction

Glossary

Acculturation level the degree to which a person from a minority group functions and interacts competently within the majority culture.

Assimilation the expectation that minorities will discard their minority culture in favor of embracing majority culture beliefs and practices.

Bicultural competence being highly competent in both majority and traditional minority cultural practices.

Columbian exchange the natural exchange of human pathogens that occurred upon first contact between old world and new world citizens during the colonial period. Diseases like small pox and tuberculosis spread rapidly through new world populations because of no natural resistance and syphilis spread rapidly through European populations for the same reason.

Community-Based Participatory Research a model that includes minority community stakeholders as equal partners with researchers in the development, implementation, and dissemination of research studies.

Construct equivalence the assumption that a particular item, factor, or variable of interest is the same conceptually across cultures and can be measured in the same way (e.g. intelligence).

Culturally bound conditions syndromes or illnesses that are defined by a group but may have no equivalent outside of that group.

Culturally specific constructs items, factors, or variables that are culturally important to understand the behavior of a group but may not exist or be relevant outside of that culture.

Discrimination treating a person differentially and in a way detrimental to the person because of bias toward that person's group membership.

Ethnic identity how a person defines her or his identity by ethnicity. The ethnic identity of a person may not agree with the ethnic identity assigned by the culture to the person, and ethnic identity can reflect multiple culturally heritages.

Empirically supported interventions interventions empirically tested by use of a randomly controlled trial experimental design and found to be significantly more effective than no intervention at all or treatment as usual.

Enculturation level the degree to which a person from a minority group functions and interacts competently within her/his minority culture.

Health disparity significantly higher prevalence of a health problem among a group than would be predicted by rates among the general population or significantly less access of care than would be predicted by the prevalence of a health problem in the group.

Implicit bias since inherent bias permeates a culture, people socialized within that culture often believe and act on biased assumptions without awareness of their biases.

Intergenerational trauma the belief by many minority people that historical incidents of oppression have created trauma so severe that it has been passed intergenerationally among the group.

Majority group the group that prescribes the appropriate belief system and behavior for a particular society. Although the majority group generally has the largest concentration of political power in a society, it is not always the numerically largest subgroup.

Microaggression often subtle insults, expressions of stereotypes, or put downs based upon group bias that may occur without the awareness of the perpetrator.

Minority groups subgroups in society that often have different assumptions about the world than the majority and sometimes have been oppressed because of these views. Minority groups can be the largest group in a society numerically, but not wield dominant political power.

Multicultural competence being highly competent in multiple cultures including the majority culture. Multicultural competence is a term recognizing that ethnic identity may involve more than two cultures.

Prejudice biased attitudes about groups of people that are insulting or injurious to a minority group member that are often expressed or acted upon automatically and without awareness.

Racism belief in biologically based group differences (often defined by skin color and phenotypes) that causes certain groups to be biologically and intellectually superior to another.

Stereotype threat minority group members experience stress that can negatively impact outcomes when engaged in a performance task that may involve a stereotyped behavior for their group.

Therapeutic alliance the ability of client and therapist to work together collaboratively on client-oriented treatment goals.

Further Reading

American Indian and Alaska Native Mental Health Research, 1989. 3 (2), [Special edition on the Barrow Alcohol Study].

American Journal of Public Health, 2003 93 (2), [Special edition on racial/ethnic bias and health].

Blume, A.W., Morera, O.F., García de la Cruz, B., 2005. Assessment of addictive behaviors in ethnic-minority populations. In: Donovan, D.M., Marlatt, G.A. (Eds.), Assessment of Addictive Behaviors, second ed.). Guilford Press, New York, pp. 49–70.

Blume, A.W., Resor, M.R., Kantin, A.V., 2009. Addiction treatment disparities among ethnic and sexual minority populations. In: Miller, P.M. (Ed.), Evidence-based Addiction Treatment. Elsevier Press, San Diego, pp. 313–325.

Steele, C.M., 1997. A threat is in the air: how stereotypes shape intellectual identity and performance. American Psychologist 52, 613–629.

Sue, D.W., Capodilupo, C.M., Torino, G.C., Bucceri, J.M., Holder, A.M.B., Nadal, K.L., Esquilin, M., 2007. Racial micro-aggressions in everyday life: implications for clinical practice. American Psychologist 62, 271–286.

Sue, D.W., Sue, D., 2003. Counseling the Culturally Diverse: Theory and Practice. John Wiley and Sons, Inc, New York.

U.S. Department of Health and Human Services, 2001. Mental Health: Culture, Race, and Ethnicity A Supplement to Mental Health: A Report of the Surgeon General. U.S. Department of Health and Human Services, Rockville, MD.

Relevant Websites

Office of Ethnic Minority Affairs, American Psychological Society, http://www.apa.org/pi/oema/.

The U.S. National Center on Minority Health and Health Disparities, http://ncmhd.nih.gov/.

U.S. National Institute on Drug Abuse Health Disparities, http://www.nida.nih.gov/about/organization/healthdisparities/.

U.S. National Institute on Alcoholism and Alcohol Abuse Strategic Plan to Address Health Disparities, http://pubs.niaaa.nih.gov/publications/HealthDisparities/Strategic.html.

CHAPTER

17

Medical Toxicology of Drugs of Abuse

Teresa Cunha-Oliveira, Ana Cristina Rego*, Félix Carvalho$, Catarina R. Oliveira**

*CNC, University of Coimbra, Coimbra, Portugal $University of Porto, Porto, Portugal

OUTLINE

INTRODUCTION

Drug addiction seriously affects public health and represents a social burden worldwide. The abuse of drugs has become a complex issue, mainly due to the development of synthesis and purification procedures that enable an increase in the effective quantities of the active compounds consumed and to the invention of the hypodermic syringe in the mid-nineteenth century, which allowed the direct injection of purified active compounds into the bloodstream. This also contributed to the increase of infections among drug addicts.

Among psychoactive drugs, alcohol (ethanol) is the most common in the world. Its harmful use is responsible for 3.8% of all global deaths. Besides alcohol, the most abused drugs in the world are cannabis (used annually by 2.9–4.3% of the world population aged 15–64), amphetamines (0.3–1.2%), cocaine (0.3–0.4%), and opiates (0.3–0.5%), as described in the World Drug Report 2010, from the United Nations Office on Drug and Crime. Although less consumed, opiates are illicit drugs that lead more people to seek treatment, due to the severe withdrawal effects and the increased risk for infections.

This chapter summarizes and compares the characteristics and toxicological properties of the main drugs of abuse, namely alcohol, amphetamines, cocaine, heroin, and cannabis.

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00017-6

CHARACTERISTICS OF PSYCHOACTIVE DRUGS

The term "drug of abuse" is usually applied to a psychotropic drug that is used in a manner that deviates from the approved medical or social patterns within a given culture at a given time.

Drugs with psychoactive effects can be divided into several groups, according to their specific actions. The most common illicit drugs of abuse are the psychostimulants (e.g. amphetamines and cocaine), depressants (e.g. alcohol (ethanol) and opiate narcotic analgesics), and hallucinogens (e.g. mescaline and lysergic acid – LSD). The properties of these groups of drugs are summarized in Table 17.1.

Some drugs of abuse induce effects that are common to more than one group. For example, ecstasy (or 3,4-methylenedioxymethamphetamine – MDMA) belongs to the class of amphetamine-type psychedelic drugs, which share stimulant and hallucinogenic effects. These drugs are also known by empathogens or entactogens, because they induce feelings of empathy and entactogeny. Another example is cannabinnoids, which share properties of all the groups described above.

Drug abuse is frequently associated with toxic effects that evolve under regular use, overdosage or the withdrawal syndrome that manifests during abstinence from the drug. It affects a number of body systems, leading to signals and symptoms of organ dysfunction, such as:

- Central nervous system (CNS) symptoms that may range from headaches and altered mental status to coma and seizures.
- Cardiovascular alterations that include changes in blood pressure, heart rate, as well as arrhythmias and organ ischemia.
- Respiratory changes that include respiratory arrest, pulmonary edema, and pneumothorax.
- Metabolic effects, including alterations in body temperature, electrolytes, and acid–base disturbances.
- Hepatic damage, from hepatitis to severe hepatotoxicity and liver failure that may require liver transplantation.
- Renal damage, with symptoms derived from decreased filtration rate to acute kidney failure.
- Reproductive consequences that may range from impaired fertility to teratogenesis, intrauterine growth retardation, premature births and neonatal syndromes, and attention deficit hyperactivity disorder (ADHD).
- Infectious complications from intravenous drug use, including viral infections such as HIV and hepatitis B, and bacterial infections such as bacterial endocarditis, osteomyelitis, and abscesses.

The clinical toxicology of drugs of abuse depends on the administration pathway, which affects its bioavailability

TABLE 17.1 Classification and Principal Effects of the Main Drugs of Abuse

	Effects			
Class	**Acute**	**Chronic**	**Withdrawal**	**Drugs**
Sedative/hypnotics	Euphoria, relaxation, CNS depression, nausea, vomiting, impaired motor function, impaired sensory function, impaired cognition	Craving, tolerance, physical dependence	Severe shaking, sweating, weakness, agitation, headache, nausea, vomiting, tachycardia, seizures	Alcohol (ethanol)
Psychostimulants	Euphoria, tachycardia, hypertension, hyperthermia, increased mental alertness, seizures	Psychosis, paranoia, reduced appetite, weight loss, heart failure, nervousness, insomnia	Severe depression (sometimes), headache	Cocaine, amphetamine and derivatives (e.g. methamphetamine, ecstasy, cathinone, mephedrone)
Opioid-type depressants	Pain relief, euphoria, drowsiness/nausea, constipation, confusion, sedation, respiratory depression and arrest, hypothermia, unconsciousness, seizures, coma, death	Depressed sexual drive, lethargy, general physical debilitation, infections, hepatitis, tolerance, addiction	Anxiety, insomnia, nausea, vomiting, diarrhea, anorexia, tachycardia, lacrimation, sweating, severe back pain, stomach cramps, muscle spasms	Opium, morphine, heroin, desomorphine
Hallucinogens	Altered states of perception and feeling	Persisting perception disorders (flashbacks)	No typical symptoms	Mescaline, LSD, psilocybin, ecstasy

(affecting the onset and extent of the psychotropic effects), the biodistribution (and therefore the exposure of target organs), and biotransformation or metabolism, which occurs mainly in the liver (affecting the nature and concentration of toxic compounds in the organism).

The reaction of the toxic compounds with their target molecules may result in their dysfunction or destruction, or in the generation of new toxic compounds. In consequence, cellular exposure to toxic compounds may result in cell dysfunction and cell death, if cell repair and adaptation are overcome. Beyond the inherent toxicity of drugs of abuse, the toxic effects may be affected by adulterants and other impurities, or even by interactions among different drugs in the frequent events of polydrug abuse.

ADULTERANTS AND CONTAMINANTS

A critical problem associated with drug abuse is the fact that the drugs available in the streets are illegally synthesized, usually under poor conditions. Deficient purification and low quality of the reagents used often leave some impurities in the final products. Frequently, adulterants are also intentionally added to the drugs to increase profit or to modulate the experienced effects.

Heroin is a semisynthetic drug, obtained from acetylation of morphine. *Street* heroin may contain different amounts of heroin and other components, depending on its origin and on the method of illicit synthesis. Usually, *street* heroin is illegally synthesized from morphine purified from opium extracts, which is often contaminated with other alkaloids. These alkaloids may also suffer synthetic acetylation during heroin manufacture. Depending on the purification procedure, *street* heroin may contain some impurities, such as morphine and 6-monoacetyl morphine (6-MAM) (heroin metabolites) or codeine and acetylcodeine. Heroin in seized samples often contains various inert diluents (starch, lactose, fructose, sucrose, mannitol, powdered milk) and active adulterants (caffeine, paracetamol, strychnine, acetylsalicylic acid, barbiturates, quinine, and amphetamines).

Street cocaine can be mixed with several diluents or adulterants, such as amphetamines, antihistamines, benzocaine, inositol, lactose, lidocaine, mannitol, opioids, phencyclidine, procaine, sugars, tetracaine, and sometimes arsenic, caffeine, quinidine, and even flour or talc.

MDMA is also frequently adulterated. Occasionally, tablets that are sold as "ecstasy" do contain drugs other than MDMA, or even none at all. Other psychoactive substances found in tablets sold as "ecstasy" included mostly other amphetamines, such as 3,4-methylenedioxyamphetamine (MDA), 3,4-methylenedioxyethylamphetamine, paramethoxyamphetamine, 2,5-dimethoxy-4-bromoamphetamine (DOB), and 4-methylthioamphetamine (4-MTA). Other compounds such as caffeine, cocaine, heroin, ketamine, LSD, aspirin, synthesis intermediaries, among other drugs, have been found in ecstasy tablets and may contribute to its toxicological outcome. The online project http://www.ecstasydata.org/ receives ecstasy tablets sent by users for testing and publishes the results in their website with the aim of helping drug users with harm-reduction, medical personnel, and researchers.

POLYDRUG USE

An important factor affecting drug toxicity and medical complications is polydrug abuse. A relatively common combination of drugs is the *speedball*, which consists in concurrent administration (by injection) of cocaine and heroin. *Speedball* has been reported to cause more rewarding effects in rats than cocaine or heroin alone. The popularity of this drug combination may be explained by the reduction of the unwanted side effects of one drug by the other, since they have different mechanisms of action, or by the enhancement of the desired effect at the reward system.

Ethanol is frequently combined with other drugs of abuse. When ethanol and cocaine are co-consumed, the euphoric effects of cocaine are enhanced. However, this combination also increases the toxic effects of both drugs, because the drugs are combined in vivo to form a very toxic metabolite – cocaethylene. This is a very lipophilic compound and is able to cross the blood–brain barrier. The effects of cocaethylene are similar to those of cocaine but the metabolite has a longer half-life, prolonging the acute effects of cocaine.

Consumption of ethanol also increases the toxic effects of MDMA, enhancing hyperthermia, hepatotoxicity, and neurotoxicity.

TOXICOLOGICAL PROPERTIES OF SELECTED DRUGS

Characteristics of drugs of abuse such as induction of positive reinforcement, dependence, and withdrawal are generally associated with certain pharmacological properties. Psychoactive drugs with rapid absorption and delivery to the CNS, high bioavailability, low protein and peripheral tissue binding, small volume of distribution, short half-life, and high free drug clearance are generally predicted to produce positive reinforcement and lead to persistent self-administration. Drugs that induce physical dependence generally have a long half-life, low free drug clearance, and must achieve high enough concentrations for sufficient time to induce the development of compensatory homeostatic changes that permanently or temporarily change the organism's

response to the drug. These homeostatic changes are responsible for the development of tolerance and sensitization to the drug, and for the withdrawal syndrome that manifests in the absence of the drug. Withdrawal symptoms are most probably manifested for psychoactive drugs with a short half-life, high free drug clearance, and that rapidly exit the CNS.

The intensity and onset of a drug's effects are determined by the rapidity of its delivery to the CNS. Drug users learn to optimize the delivery of the drug to the brain and to maximize the bioavailability of the drug by adapting the methods and routes of administration. The most rapid CNS delivery is achieved by inhalation, due to the direct access of pulmonary blood to the brain. Smoking is a very effective route of administration, but requires highly volatile forms of the drugs, to assure resistance to degradation at the temperatures produced by burning. Intravenous administration provides the highest bioavailability but is associated with severe health complications, whereas oral administration is generally more convenient, but is associated with lower bioavailability and slower delivery to the brain.

Central effects of drugs of abuse are due to the interference of these drugs with the molecular and cellular pathways involving neuronal active endogenous compounds, including monoamine and other types of neurotransmitters, endocannabinoids and endorphins, taking advantage of structural and bioisostere similarities (Fig. 17.1). The molecular targets of these xenobiotics are transporters or receptors that mediate the physiological actions of those endogenous compounds, activating specific intracellular signaling pathways. However, drugs of abuse do not completely mimic the action of the endogenous compounds because the molecular machinery involved in their removal from the synapse is frequently inefficient for these xenobiotics, which implies the interference with neuronal activity for longer periods of time.

In the next sections, we discuss the toxicological properties of the most common drugs of abuse, the main medical complications found in drug abusers and the biomarkers of abuse. The data presented are mainly based on literature referred at the Further Reading section.

Alcohol (Ethanol)

Alcohol (ethanol) has long been used by mankind for social, medical, cultural, and religious purposes. For most users, alcohol consumption does not impair physical or mental health. However, acute or chronic alcohol intoxication has negative individual and social consequences. Ethanol is a straight-chain alcohol (Fig. 17.1) produced by fermentation of sugars present in agricultural products. It is present in alcoholic drinks, namely in beer (3–6% w/v), wine (9–12%), spirits (32–40%), or cocktails (15–25%). Alcohol is generally used to obtain euphoria and relaxation. The effective dose in humans is around 22–40 g and the effects generally persist for 1.5–3 h. The harmful use of alcohol is generally associated with chronic or binge drinking.

Routes of Exposure

Alcohol is generally used by the oral route to obtain its psychoactive effects. However, it may be used as an antiseptic for medical purposes, applied topically in the skin.

Pharmacokinetics

- Absorption

When taken orally, alcohol is rapidly absorbed in the small intestine into the bloodstream. The presence of food in the stomach may delay gastric emptying and slow absorption. The oral bioavailability of alcohol is generally higher than 80%. Peak blood concentrations typically occur between 30 and 90 min after ingestion.

- Biodistribution

Alcohol has low molecular weight, mixes well with water and is only weakly charged, easily crossing biological membranes and the blood–brain barrier. It has low-protein binding and a volume of distribution around 0.55 l kg^{-1}. On entering the bloodstream, alcohol is distributed throughout the body, mainly affecting the brain, liver, and kidneys.

- Plasma half-life

At high alcohol concentrations, the plasma half-life is about 4–5 h, whereas at low alcohol concentrations the clearance is slower and the plasma half-life increases.

- Metabolism

Alcohol is mainly metabolized in the liver by alcohol dehydrogenase into acetaldehyde. Alcohol dehydrogenase may also be present in the stomach and small intestine. Catalase and cytochrome P450 (CYP2E1) may also contribute to alcohol metabolism into acetaldehyde, and hepatic CYP2E1 expression is 5- to 10-fold increased in chronic alcohol users. Acetaldehyde is then converted into acetate by aldehyde dehydrogenases (these enzymes are found in many tissues of the body but are at the highest concentration in the liver). Acetate combines with coenzyme A, to generate acetyl-coenzyme A, which may enter metabolic pathways. A small percentage of ethanol is conjugated to give ethyl glucuronide and ethyl sulfate, which may be helpful as biomarkers of alcohol abuse.

- Excretion

About 5–10% of ingested alcohol is excreted unchanged in urine, breath, and sweat. Alcohol or its

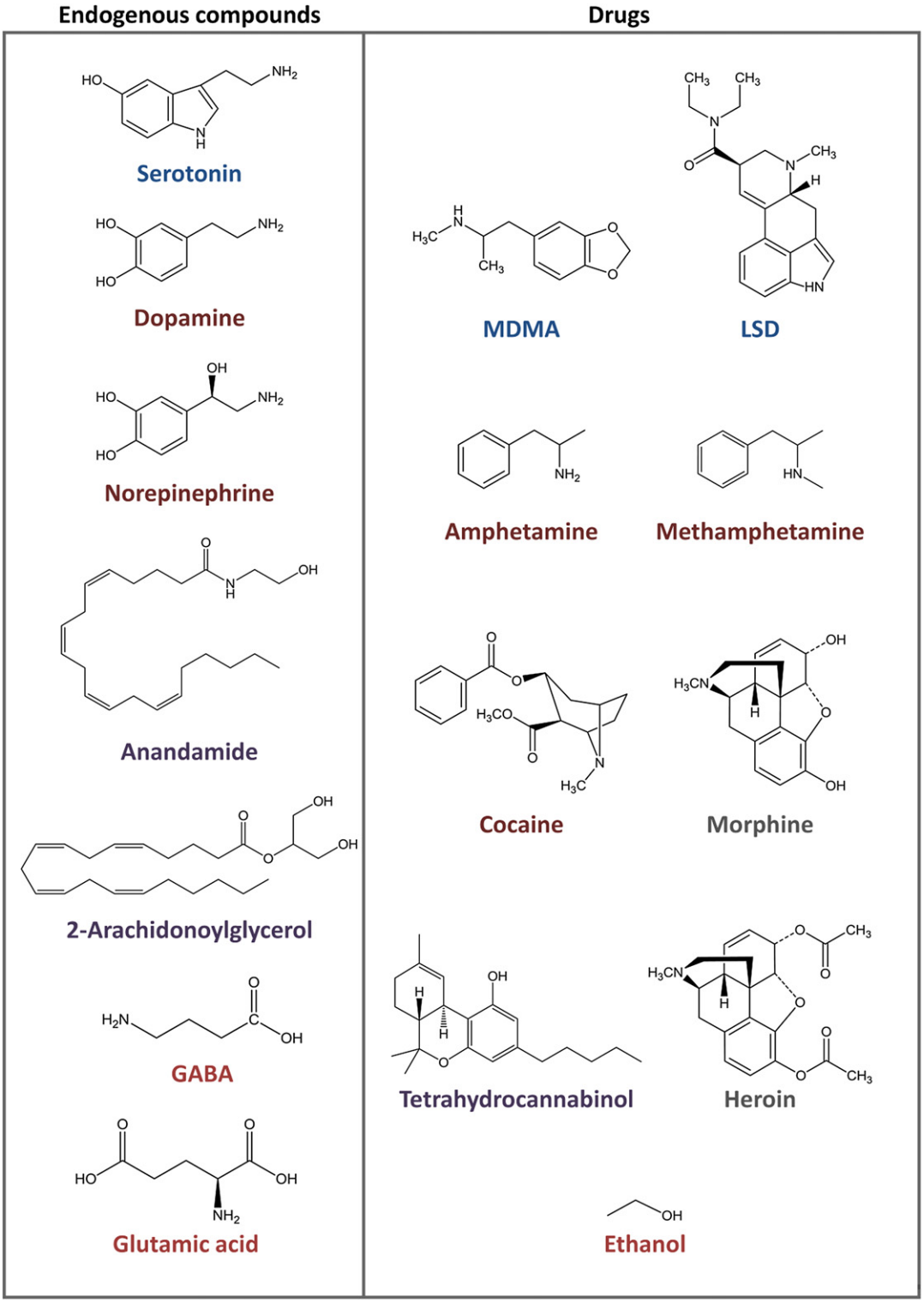


FIGURE 17.1 Chemical structures of some drugs of abuse and neuronal active endogenous compounds. The structures of MDMA and LSD resemble the structure of serotonin and thus interfere with serotonergic systems. Amphetamine, methamphetamine, and cocaine are structurally similar to dopamine and norepinephrine, affecting primarily the systems involving these monoamines. Tetrahydrocannabinol has similarities with the endogenous cannabinoids anandamide and 2-arachidonylglycerol and interferes with the receptors for these compounds. Morphine and heroin present chemical similarities to the active sites of endogenous opioid polypeptides such as enkephalins and endorphins (polypeptidic structures not shown here for simplicity). Ethanol shares structural similarities with the neurotransmitters GABA and glutamate and interferes with receptors for these amino acids.

metabolites are generally detected in urine up to 96 h. Alcohol use is generally tested by breath analysis, which is well correlated with blood alcohol concentration.

Pharmacology and Toxicology

- Pharmacodynamics/Mode of action

Although alcohol has long been believed to act nonspecifically by disordering lipids in cell membranes, it is now acknowledged that at physiologically relevant concentrations (5–20 mmol l^{-1}), alcohol interferes with neural activity by acting directly on neurotransmitter-gated ion channels. Ethanol enhances GABAergic neurotransmission by altering the conformation of inhibitory $GABA_A$ receptors and inhibits the excitatory N-methyl-D-aspartate (NMDA) receptors, acutely depressing neural activity, thus explaining the sedative effect of alcohol.

At the reward pathway, ethanol acts on $GABA_A$ receptors present on inhibitory neurons in the ventral tegmental area, and induces the release of opioid neuropeptides, leading to the disinhibition of dopamine release in the nucleus accumbens. Ethanol may also act in the nucleus accumbens, possibly by inhibiting NMDA receptors in corticostriatal synapses.

Long-term alcohol use leads to neuroadaptations at the ion channel sensitivity (subunit composition) or number. When alcohol is no longer present, the decrease in inhibitory and increase in excitatory receptor functions become unmasked, leading to the withdrawal syndrome, where neurons are in a hyperexcitable state.

- Toxicity

Ethanol is a general CNS depressant, inducing a degree of sedation dependent on the blood concentration achieved. On drinking a moderate amount of ethanol, users may experience a stimulating phase, due to the depression of the brain mechanisms that control behavior. At low-blood alcohol concentrations (0.01–0.1% w/v), the main brain areas affected are at first the cerebral cortex and then the forebrain, associated with the feelings of relaxation, well-being, loss of inhibition, pleasure, and emotional arousal. From concentrations of 0.01–0.3%, the cerebellum and the brain stem also become affected, leading to mood swings, anger, sadness, aggression, and depression. From 0.31% to higher blood alcohol concentrations, the entire brain becomes affected, leading to unconsciousness, coma, and possibly death. Accordingly, the level of alcohol-induced impairment is also dependent on blood alcohol concentration. At first, alertness is affected, followed by judgment, motor coordination, visual tracking, balance, temperature regulation, bladder control, breathing, and heart rate.

Chronic alcohol use induces liver damage, including fat accumulation, alcoholic hepatitis, and cirrhosis. Chronic alcohol intake also affects digestive functions, leading to gastritis and pancreatitis; cardiovascular function, by inducing cardiomyopathy, arrhythmias, and hypertension; brain damage, head, neck, and esophageal cancers; and also affects immune and endocrine systems. Undernutrition may also be observed in chronic alcohol users, particularly involving vitamin deficiencies.

Tolerance to alcohol develops rapidly, due to neuroadaptations and induction of metabolic enzymes. These adaptations underlie the withdrawal syndrome, which highly contributes to the burden of alcoholism. Alcohol withdrawal is characterized by the symptoms of autonomic nervous system hyperactivity. The initial symptoms are usually mild and include anxiety, insomnia, and tremors, beginning within about 3–6 h of last alcohol intake and usually lasting about 1–3 days. In 5–10% of patients, severe convulsions may also occur in the first 2 days of abstinence. About 10% of alcoholics may develop more severe withdrawal symptoms involving autonomic nervous system hyperactivity, including increases in blood pressure, pulse, breathing and heart rates, and body temperature. Excessive sweating and tremors generally occur. In extreme cases, severe alcohol withdrawal may be complicated by the presence of delirium tremens," which usually manifests within 48–72 h of abstinence. This condition involves agitation, confusion, disorientation, delusions, and vivid hallucinations, and may persist up to 96 h of drink cessation.

Fetal alcohol syndrome (FAS) is an important complication of alcohol abuse by pregnant women. Children with FAS present developmental anomalies such as deficits in the formation of the CNS and restricted physical growth, which may lead to cognitive, behavioral, emotional, and social deficits.

Blood concentrations found in ethanol-related deaths are in the range 2.2–5.0 g l^{-1} and a typical lethal dose varies between 276 and 455 g.

The median lethal dose (LD50) in rats ranges between 5.6 and 10 g kg^{-1} when consumed orally. In mice, the LD50 (oral) was reported to be 3.45 g kg^{-1}.

Methamphetamine (and Other Amphetamines)

Amphetamine was first synthesized in 1887 by Lazar Edeleanu at the University of Berlin. This drug is a synthetic derivative of the plant alkaloid ephedrine, extracted from plants in the genus *Ephedra*. *Ephedra sinica*, also known as Ma Huang, has been used in traditional Chinese medicine for 5000 years to treat several diseases, such as asthma and common cold. Amphetamines are illegally used to increase alertness, to relief fatigue, control weight, and for their intense euphoric effects. Amphetamines are still used in medical practice

to treat narcolepsy and ADHD and have been used as energy boosters by athletes, soldiers, fighter aircraft pilots, and long distance truck drivers.

Methamphetamine is a common amphetamine derivative, more potent than the parent compound. Methamphetamine hydrochloride is presented as a white to light-brown crystalline powder, or crystals that resemble ice, whereas methamphetamine base is a liquid. Methamphetamine is usually available in high purity forms, ranging from 60 to 90% purity. A typical dose ranges from 50 to 2000 mg day^{-1}, but in chronic binge users it may reach 5000 mg day^{-1}.

Another popular amphetamine derivative is MDMA, a ring-substituted amphetamine derivative with mild hallucinogenic properties. It was first synthesized and patented in 1912 by the German pharmaceutical company Merck under the name of "methylsafrylamin," as a precursor for therapeutically active compounds. In 1976, MDMA was used for the first time in the clinics as an adjuvant to psychiatric treatment, to increase patient self-esteem and facilitate therapeutic communication, which continued until the early 1980s, when MDMA was classified as a schedule one drug due to its high abuse potential, lack of clinical application, lack of accepted safety for use under medical supervision, and evidence that it could be neurotoxic. Also at the early 1980s, it became popular in the streets as a recreational drug and is still highly used nowadays, especially in dance parties (raves). MDMA is a white, tan or brown powder, primarily available in tablet form. The typical content of MDMA per tablet has been reported to range from 2 to 130 mg, although the average is between 30 and 80 mg. The typical pattern of MDMA use ranges from 1 to 2 tablets in a single episode, though binge administration of ecstasy tablets is also frequent among users.

Routes of Exposure

Methamphetamine users generally begin with intranasal or oral use and may progress to intravenous use, and occasionally smoking. "Ecstasy" is almost exclusively sold and consumed orally in the form of tablets (rarely capsules), which frequently contain symbols (logos) and are colored.

Pharmacokinetics

- Absorption

Amphetamines are rapidly absorbed after oral ingestion, with peak plasma levels occurring within 2.6–3.6 h. Peak plasma concentrations may range from 0.01 to 2.5 mg ml^{-1} for methamphetamine or 0.02–0.44 mg ml^{-1} for MDMA. The effects of methamphetamine usually persist for 4–8 h, but residual effects may last up to 12 h. MDMA effects may persist for 2–3 h.

- Biodistribution

Amphetamines concentrate in the liver, kidney, lungs, cerebrospinal fluid, and brain. They are highly lipid soluble and readily cross the blood–brain barrier. Amphetamines are weak bases with high p*K*a values between 9.4 and 10.1, low molecular weight, low protein binding (around 20%), and high volume of distribution (3.5–7 l kg).

- Biological half-life

The biologic half-life of orally administered methamphetamine is 10.1 h in average, ranging from 6.4 to 15 h, while that of MDMA is found to be in the range of 6–9 h, depending on the dose.

- Metabolism

The phase I metabolism of methamphetamine by CYP2D6 generates two pharmacologically active metabolites, amphetamine and 4-hydroxymethamphetamine. The major metabolic pathway for amphetamine involves aromatic hydroxylation by CYP2D6 to 4-hydroxyamphetamine, which is psychoactive, and deamination to phenylacetone. This compound is subsequently oxidized to benzoic acid and excreted as glucuronide or glycine (hippuric acid) conjugate. Smaller amounts of amphetamine are converted to norephedrine by oxidation. Hydroxylation of norephedrine produces an active metabolite, 4-hydroxynorephedrine, which is psychoactive.

MDMA is also a substrate for CYP2D6. The major pathways of MDMA metabolism are N-demethylation, O-demethylation, and deamination. MDMA is converted to the catechol, 3,4-dihydroxymethamphetamine (DHMA) and the N-demethylated psychoactive product, 3,4-methylenedioxyamphetamine, MDA, by CYP2D6, but other enzymes may also contribute (e.g. 1A2, 2B6, and 3A4). MDA is further metabolized to the catechol intermediate, 3,4-dihydroxyamphetamine (DHA). DHMA and DHA can undergo oxidation to the corresponding ortho-quinones, which can form adducts with glutathione and other thiol-containing compounds.

- Excretion

An oral dose of 30–54% of methamphetamine is excreted in urine as unchanged methamphetamine, and 10–23% as unchanged amphetamine. After intravenous use, 45% is excreted as methamphetamine and 7% as amphetamine. However, the amount of urinary excretion and metabolism is highly pH dependent, with alkaline urine significantly increasing the drug half-life. Detection time of amphetamine in urine is usually 1–4 days. Methamphetamine may be detected in urine 3–5 days after the last use. The urinary recovery of MDMA is approximately 60%, independently of the dose

administrated. DHMA is the main metabolite found in urine (>20%), with less than 2% of the dose excreted as MDA. MDMA or its metabolites may be detected in urine 1–5 days after the last use.

Pharmacology and Toxicology

- Pharmacodynamics/Mode of action

Methamphetamine increases synaptic levels of the monoamine neurotransmitters dopamine, serotonin (5-HT), and norepinephrine, and has α- and β-adrenergic agonist effects. Due to its structural similarity with dopamine (Fig. 17.1), amphetamine is a substrate for the dopamine transporter (DAT). Amphetamine also interferes with the vesicular monoamine transporter-2 (VMAT-2) function, impairing the active transport of the monoamines into synaptic vesicles, where they are stored. In addition, amphetamines can also deplete vesicular biogenic amine content by disrupting the pH gradient via a weak base effect that drives the transporter. Cytosolic dopamine is then released to the extracellular space via reverse transport through DAT. Amphetamine also inhibits dopamine synthesis by inhibiting tyrosine hydroxylase and may also slowdown catecholamine metabolism by acutely inhibiting monoamine oxidase.

MDMA is similar in structure and effects to methamphetamine (Fig. 17.1), but has significantly less CNS stimulant properties. MDMA has a high affinity for 5-HT2 receptors and may cause acute depletion of presynaptic 5-HT, depression of 5-HT synthesis, and retrograde destruction of 5-HT neurons. MDMA easily diffuses across the cell membranes and lipid layers and may be specifically accumulated inside serotonergic neurons through the serotonin transporter (SERT). MDMA also increases the levels of norepinephrine and dopamine. MDMA hallucinogenic properties depend on the stimulation of 5-HT2A-receptors, mainly in the pyramidal neurons of the neocortex.

Increase in synaptic dopamine in the brain reward pathway is associated with feelings of pleasure induced by amphetamines and other drugs of abuse.

- Toxicity

The increase in dopamine levels induced by amphetamines leads to an increase in its oxidative metabolism, which generates free radicals that may induce cytotoxicity. Dopamine mediates locomotor stimulation, psychosis, and perception disturbances, whereas changes in norepinephrine levels are associated with alerting, anorectic, locomotor, and sympathomimetic effects and 5-HT is responsible for delusions and psychosis. Toxicity of methamphetamine may lead to renal and liver failure, hyperthermia, cardiac arrhythmias, heart attack, cerebrovascular hemorrhages, stroke, seizures, and death.

The effects of methamphetamine are similar to those of cocaine, but start slower and last longer. In most methamphetamine-related deaths, blood concentrations found are in the range 1–43 mg l^{-1}.

D-amphetamine has similar effects to methamphetamine, but is less potent. LD50 of amphetamine was reported as 55 mg kg^{-1} (oral) and 180 mg kg^{-1} (s.c) in rats and 24.2 mg kg^{-1} (oral) in mice. LD50 for methamphetamine was reported as 70 mg kg^{-1} (ip) in rats, 43 mg kg^{-1} (ip) in mice, and 10 mg kg^{-1} (oral) in dogs. The lethal dose of methamphetamine in humans is usually within 140–1650 mg.

MDMA associated fatalities have been reported with blood levels of 0.04–8.5 mg l^{-1}. LD50 for MDMA was determined as 97 mg kg^{-1} (ip) in mice, 49 mg kg^{-1} (ip) and 160 mg kg^{-1} (oral) in rats, and 26–98 mg kg^{-1} (ip) in guinea pigs. The lethal dose of MDMA in humans is in the range of 150–1250 mg.

MDMA neurotoxicity is the most widely studied toxic effect and potentially the most significant long-term effect of this drug. Other complications of acute MDMA use include hyperthermia, arrhythmias and cardiovascular collapse, liver failure, renal failure, and hyponatremia. Another severe consequence is rhabdomyolysis, which is characterized by the breakdown of muscle fibers that result in the release of their myoglobin contents into the bloodstream, contributing to kidney damage.

The MDMA metabolite, MDA, induces higher levels of stereotypic behavior and is more neurotoxic than the parent drug. MDA destroys 5-HT-producing neurons, which regulate aggression, mood, sexual activity, sleep, and sensitivity to pain.

Cocaine

Cocaine was first isolated in 1855 by the German chemist Friedrich Gaedcke. This alkaloid is extracted from the plant *Erythroxylum coca*, which is cultivated in the South American countries Bolivia, Colombia, and Peru. The natives of these countries chew the coca leaves in magical ceremonies and initiation rites. Cocaine may be processed in water-soluble or -insoluble forms. Water-soluble forms include cocaine sulfate and cocaine hydrochloride. In medicine, cocaine is used as a topical local anesthetic for ear, nose, and throat surgery.

Cocaine hydrochloride is presented as a shiny white to light-brown crystalline powder, whereas cocaine base is generally a white to beige waxy solid. Cocaine is used recreationally to increase alertness, relief fatigue, and increase self-confidence, and is abused for its intense euphoric effects. Purity of cocaine hydrochloride ranges from 20 to 95%, and crack cocaine is generally 20–80% pure. Cocaine is often "cut" with sugars, other CNS stimulants, and local anesthetics. Common doses range from 10 to 120 mg.

Routes of Exposure

Cocaine sulfate and cocaine hydrochloride are used by oral, sublingual, intranasal, and intravenous routes, whereas coca leaves may be chewed. Drug smugglers, known as "mules" or "body packers," may swallow packages of cocaine, which may leak or rupture and cause massive intoxication.

Water-insoluble forms such as free base cocaine or crack are usually smoked. Crack cocaine is abused by inhaling the vapor from cigarettes (usually mixed with tobacco or marijuana) or after heating the drug in a glass pipe. Most drug abusers use cocaine by the nasal route. Cocaine hydrochloride can be "sniffed" or "snorted" in "lines" on a flat surface. This route leads to pulmonary complications.

Some drug abusers inject cocaine hydrochloride subcutaneously, intramuscularly, or intravenously, alone or with heroin ("speedball") or with other drugs

Cocaine can also be administered rectally, vaginally, and urethrally.

For clinical purposes, cocaine is used topically to take advantage of its local anesthetic effects.

Pharmacokinetics

- Absorption

Cocaine is rapidly absorbed following smoking, snorting, and intravenous administration.

Bioavailability is about 93.7% after intranasal use and 70% on smoking.

Injecting cocaine produces an effect within 15–30 s. After smoking crack or snorting cocaine the effects are almost immediate. In fact, through this way, cocaine enters the pulmonary circulation and reaches the cerebral circulation within 6 s, eliciting a rapid, short, but very intense euphoric effect. The effects of crack typically last 5–15 min, whereas after snorting the effects may last 15–30 min.

When orally ingested, the effects of cocaine begin to be observed in about 1 h and may persist for 1–2 h.

Typical blood concentrations after a single use are in the range of 0.2–0.4 mg ml^{-1}, but tolerant individuals may present up to 5 mg l^{-1}.

- Biodistribution

Cocaine is distributed within all body tissues, and crosses the blood–brain barrier. In large, repeated doses, it is probably accumulated in the CNS and in adipose tissue, due to its lipid solubility. Cocaine is found 91% bound to proteins and its volume of distribution varies between 1 and 3 l kg^{-1}. Cocaine crosses the placenta by simple diffusion and may accumulate in the fetus after repeated use.

- Biological half-life

Cocaine half-life is about 1 h, varying from about 0.6 h on smoking, 0.8 h after oral administration, 1.25 h after nasal administration, and 0.7–0.9 h after parenteral administration.

- Metabolism

Cocaine metabolism takes place mainly in the liver, within 2 h of administration. The rate of metabolism varies according to plasma concentration. There are three main routes of biotransformation: The major route is hydrolysis of cocaine by hepatic and plasma esterases, with loss of a benzoyl group originating ecgonine methyl ester (EME). The secondary route is spontaneous hydrolysis, which leads to benzoylecgonine (BE) by demethylation. Both EME and BE are then converted into ecgonine. A minor route is N-demethylation of cocaine by CYP3A4, leading to the active metabolite norcocaine, which crosses the blood–brain barrier.

Anhydroecgonine methyl ester can be produced when the drug is consumed in the free base form (as a result of thermal degradation of smoked "crack"). In the presence of alcohol another active metabolite, cocaethylene, is formed, which is more toxic than cocaine itself.

- Excretion

Unchanged cocaine is recovered at less than 2% in urine, although higher proportions may be seen in acidic urine; 26–39% of cocaine is recovered as BE and 18–22% as EME. After 4 h of use, most of the drug is eliminated from plasma. Cocaine metabolites persist in urine at detectable concentrations from 2 to 4 days of abstinence, but after chronic use they may be present for up to 10 days after the last use.

Pharmacology and Toxicology

- Pharmacodynamics (mode of action)

The main targets of cocaine are the CNS and cardiovascular system.

Cocaine interferes with the reuptake of monoamine transmitters, particularly dopamine, a neurotransmitter associated with pleasure and movement. Cocaine binds to the DAT blocking its function, which leads to increased extracellular dopamine and results in chronic stimulation of postsynaptic dopamine receptors, resulting in the euphoric "rush." Dopamine levels then fall, resulting in the dysphoric "crash." Cocaine also interferes with the uptake of norepinephrine and 5-HT, leading to accumulation of these neurotransmitters at postsynaptic receptors. Cocaine also acts as a local anesthetic, because it reversibly blocks the initiation and conduction of the nerve impulse, by binding to voltage-gated sodium channels.

Cocaine also increases catecholamine concentrations in the blood, leading to excessive stimulation of peripheral α- and β-adrenoreceptors.

- Toxicity

The neurotoxic actions of cocaine involve several brain areas and different mechanisms of action. Euphoria, confusion, agitation, and hallucination result from an increase in dopamine activity in the limbic system. Cortical effects lead to pressure of speech, excitation, and a reduced feeling of fatigue. Stimulation of lower centers leads to tremor and tonic–clonic convulsions. Brain stem effects lead to stimulation and then depression of the respiratory vasomotor and vomiting centers. Cocaine may induce hyperthermia due to increase in muscular activity and by a direct effect on thermal regulatory centers.

At low doses, cocaine induces vagal stimulation with bradycardia, whereas at moderate doses, adrenergic stimulation leads to a rapid increase in cardiac output, myocardial oxygen consumption, and blood pressure, then followed by a decrease. This may result in increased risk of myocardial infarction and spontaneous cerebral hemorrhage. At very high doses, a direct toxic effect of cocaine on the myocardium may result in cardiac arrest.

Cocaine abusers may present rhabdomyolysis, probably due to a direct effect of cocaine on muscle and muscle metabolism, tissue ischemia, or due to the effects of other drugs taken with cocaine, such as alcohol and heroin.

Prenatal brain toxicity constitutes another serious negative effect of cocaine, leading to structural, metabolic, and functional brain abnormalities.

Lethal doses of cocaine in humans are estimated at 20–2000 mg. However, cocaine addicts can tolerate doses up to 5 g day^{-1}. Toxic effects can be manifested with plasma concentrations of 0.50 mg l^{-1} or more and death has been reported with concentrations of 1–20 mg l^{-1}.

The LD50 of cocaine was determined as 17.5 mg kg^{-1} (iv) in rats, 91 mg kg^{-1} (ip) in mice, and 21 mg kg^{-1} (iv) in dogs.

Heroin and Morphine

Morphine and heroin are derived from opium, which is extracted from the opium poppy *Papaver somniferum*. There are reports of cultivation of this plant in the Mesopotamia since 3400 BC. Opium contains about 40 alkaloids that make up 10–20% of total opium substances. The most abundant opium alkaloids are morphine (8–17%), codeine (0.7–5%), thebaine (0.1–1.5%), papaverine (0.5–1.5%), and noscapine (or narcotine, 1–10%). Morphine is purified from opium extracts and converted into heroin by acetylation. Heroin is more lipid soluble than morphine and is easily transported across the blood–brain barrier, being two to four times more potent than morphine. Heroin was first synthesized in 1874 by Charles Alder Wright in England, but it was only discovered by the medical community when it was independently resynthesized, 23 years later, by Felix Hoffmann, who worked for Bayer. The use of heroin was thought to be a potential cure for morphine addiction until it was found that heroin is converted into morphine, when metabolized in the liver.

Morphine and heroin are generally white, crystalline powders. Illicit heroin may vary in color from white to dark-brown due to impurities or may appear as a black tar-like material. Depending on the demographic region, the street purity of heroin can range from 20 to 90%. Heroin may be "cut" with inert or toxic adulterants such as sugar, starch, powdered milk, quinine, and ketamine.

Heroin is often mixed with stimulants, such as methamphetamine or cocaine ("speedball"), and injected. It may also be coadministered with MDMA or crack cocaine.

Daily heroin doses may range between 5 and 1500 mg, with an average daily dose of about 300–500 mg, which may be divided by two to four daily injections.

Routes of Exposure

Morphine may be used by oral, intramuscular, intravenous, subcutaneous, rectal, epidural, and intrathecal administration. Heroin may be smoked (referred to in street jargon as "chasing the dragon"), snorted or injected intravenously ("mainlining"), and subcutaneously ("skin popping"). Black tar heroin is typically dissolved, diluted, and injected, while higher purity heroin is often snorted or smoked.

Pharmacokinetics

- Absorption

The absorption of heroin is 1.5 times higher than that of morphine, and it is 2–4 times more potent, and 200 times more soluble.

Tolerance makes interpretation of blood or plasma morphine concentrations extremely difficult. Postmortem analyses after heroin overdoses found blood morphine concentrations between 0.1 and 2.8 mg ml^{-1}.

- Biodistribution

The oral bioavailability of morphine is 20–40%, and 12–35% is bound in plasma, mainly to albumin, with approximately 5% bound to γ-globulin and 5% to α1-acid-glycoprotein. The blood/plasma concentration ratio is about 1.02 in healthy individuals. The bioavailability of smoked heroin is about 44–61%.

Morphine is relatively hydrophilic and therefore distributes slowly into tissues. The volume of distribution of morphine is generally within 1–5 l kg^{-1}. Heroin and 6-MAM cross the blood–brain barrier more easily than morphine. Morphine is transported into the brain by P-glycoprotein, present in brain capillary endothelium, and accumulates especially in the hippocampus where there is also a high concentration of opioid receptors.

- Biological half-life

Heroin has an extremely short half-life of 2–6 min. The half-lives of 6-MAM and morphine are 6–25 min and 1.5–7 h, respectively.

- Metabolism

Heroin is rapidly metabolized to 6-MAM and morphine. Heroin and 6-MAM are more lipid soluble than morphine and thus enter the brain more readily. Morphine is primarily glucuroconjugated at positions 3 and 6, to form morphine-3-glucuronide (M3G) and morphine-6-glucuronide (M6G), respectively. A small amount (5%) is demethylated to normorphine by CYP3A4 and to a lower extent by CYP2C8. M6G is an active metabolite with a higher potency than morphine. The half-life of M6G is 4 ± 1.5 h. About 90% of a single morphine dose is eliminated in urine in 72 h, 75% as M3G and less than 10% as unchanged morphine.

- Excretion

Positive morphine results in urine generally indicate use within the last 2–3 days, or longer after prolonged use. Detection of 6-MAM in urine is indicative of heroin use. High concentrations may indicate chronic use of the drug.

Pharmacology and Toxicology

- Pharmacodynamics (mode of action)

Morphine produces its major effects on the CNS primarily through μ-Receptors, and also at κ- and δ-receptors. μ-Receptors are almost always located presynaptically. Interaction of opioids with μ-opioid receptors located in inhibitory GABAergic interneurons in the reward pathway leads to inhibition of these neurons, resulting in disinhibition of dopaminergic neurons and increased synaptic dopamine concentrations associated with reward and repetitive drug use.

The brain region that contains the greatest concentration of μ-opioid receptors is the periaqueductal gray, but they are also found in the hippocampus, the superficial dorsal horn of the spinal cord, the external plexiform layer of the olfactory bulb, the nucleus accumbens (involved in reward and addiction), in some parts of the cerebral cortex, and in the amygdala.

μ1-Receptors are involved in pain modulation, analgesia, respiratory depression, miosis, euphoria, and decreased gastrointestinal activity; μ2-receptors are involved in respiratory depression, drowsiness, nausea, and mental clouding; κ-receptors are involved in analgesia, diuresis, sedation, dysphoria, mild respiratory depression, and miosis; and δ-receptors are involved in analgesia, dysphoria, delusions, and hallucinations.

Heroin has little affinity for opiate receptors. It behaves as a highly lipophilic transporter of morphine and induces more rapid and more intense CNS effects. Most of its pharmacology resides in its metabolism to the active metabolites, 6-MAM, morphine, and M6G.

Depending on morphine dose and the route of administration, effects begin within 5–60 min and may last 4–6 h. Following heroin use, the intense euphoria generally lasts from 45 s to several minutes. Peak effects may last 1–2 h, and the overall effects disappear in 3–5 h.

- Toxicity

The toxic and lethal doses depend greatly on the individual's tolerance to the drug, and thus the usual dose for an addict may be dangerous for the same individual after several days of abstinence, due to the rapid decrease in tolerance. A dose of 20 mg heroin may be lethal in non-tolerant subjects, whereas addicts may tolerate doses 10 times larger. Fatalities have been observed after a dose of 12 mg, resulting from respiratory depression.

Plasma concentration of morphine after lethal overdosage of heroin is generally in the range 0.1–2.8 mg l^{-1}, and the lethal dose is within 12–180 mg.

LD50 of heroin is around 21.8 mg kg^{-1} (iv) in mice and 23 mg kg^{-1} in rats, whereas for morphine the LD50 in mice is 226–318 mg kg^{-1} (iv).

Chronic heroin addicts frequently suffer from rhabdomyolysis, probably due to compression of muscle during prolonged immobilization, aggravated by the occlusion of vascular supply. Renal damage may progress to terminal renal insufficiency. Respiratory and cutaneous complications are also observed, probably due to immune deficiency, which may be related with a reduction in lymphocyte proliferation, spontaneous cytolytic activity, phagocytosis, and interferon production. Splenomegaly due to antigenic stimulation has been also described in heroin addicts.

Cannabis

Cannabis, also known as marijuana, is a term applied to preparations of the cannabis plant, especially of *Cannabis sativa*, intended for use as a psychoactive drug or for medicinal purposes. Cannabis is the most widely used illicit psychotropic drug in the world. The first descriptions of medical and toxic properties of the plant were part of the ancient Chinese herbal Pen-ts'ao, dating

TABLE 17.2 Characteristics and Pharmacokinetic Properties of Selected Drugs of Abuse

Drug	Ethanol	Amph	Meth	MDMA	COC	HER	MOR	THC
Class	Sedative/hypnotic, CNS depressant	CNS stimulant, sympathomimetic, appetite suppressant	CNS stimulant, sympathomimetic, appetite suppressant	Mild CNS stimulant, empathogen, entactogen, mild hallucinogen and psychedelic, appetite suppressant	CNS stimulant, local anesthetic	Narcotic analgesic	Narcotic analgesic	Cannabis/ marijuana
Main targets	$GABA_A$ receptors, NMDA receptors	Monoamine terminals	Monoamine terminals	Monoamine terminals (5-HT)	Monoamine terminals DAT	Pro-drug (μ-opioid receptors)	μ-opioid receptors	Cannabinoid receptors, CB1 and CB2
Chemical formula	C_2H_5OH	$C_9H_{13}N$	$C_{10}H_{15}N$	$C_{11}H_{15}NO_2$	$C_{21}H_{30}O_2$	$C_{21}H_{23}NO_5$	$C_{17}H_{21}NO_4$	$C_{20}H_{25}N_{3O}$
Molecular weight	46.07	135.21	149.24	193.25	303.35	369.42	285.54	314.5
Purity	Beer: 3–6% Wine: 9–12% Spirits: 32–40% Cocktails: 15–25%	Mainly pharmaceutical forms	60–90%	0–100%	HCl: 20–95% Crack: 20–80%	Samples from: Asian SW: 60–90% Middle East: 30–80% Asian SE: ~20%	Mainly pharmaceutical forms	Marijuana: 1–5% Hashish: 5–15% Hash oil: >20% Sinsemilla: up to 17%
Routes of administration	Oral	Oral	Intranasal; oral; intravenous; smoked	Oral	Topically; chewed (leaves); smoked (crack); intranasal	Smoked, snorted, intravenous, subcutaneous (skin popping)	Oral; intramuscular; IV; rectal; epidural, intrathecal	Smoked, oral
Common dose (mg)	22 000–40 000 (10–15 g/serving)	10–100	50–2000	50–2500 (average 30–80)	10–120	5–1500 (average 300–500)	60–120	5–25
Bioavailability	80%	Oral: 67.2% Smoked: 90.3%	Oral: 67.2% Smoked: 90.3%	–	Intranasal: 93.7% Smoked: 70%	Smoked: 44–61%	Oral: 20–40%	Smoked: 8–50% Oral: 4–12%
Usual blood levels ($mg\ l^{-1}$)	100–4000	–	0.01–2.5; (average 0.6)	0.02–0.44	Single dose: 0.2–0.4; repeated: up to 5	n.d.	0.20–2.3	0.1–0.2
Peak	30–90 min	–	Oral: 2.6–3.6 h; shortly after injection; few minutes after smoking	20–30 min	15–30 s	<10 min	Oral: 60 min IV: 5 min	10–30 min

Duration	1.5−3 h	−	4−8 h	2−3 h	Oral: 1−2 h; Smoked: 5−15 min Intranasal: 15 −30 min	1−2 h	4−6 h	3−5 h
Protein binding	Low	~20%	~20%	~20%	91%	0%	12−35%	95−99%
Vd ($l\,kg^{-1}$)	0.55	3.5−6.1	5−6	6−7	1−3	−	1−5	4−14
$T_{1/2}$		4−30 h	6.4−15 h (average 10.1 h)	6−9 h	Smoked: 0.6 h Oral: 0.8 h; intranasal 1.25 h; parenteral 0.7−0.9 h 4.5 h BE 3.1 h EME	2−6 min; 6-MAM: 6 −25 min	1.5−7 h; M6G: 4 ± 1.5 h	1−5 days
Metabolizing enzymes (metabolite formed)	Alcohol dehydrogenases CYP2E1 Catalase (Acetaldehyde)	CYP2D6 (4-OH-Amph)	CYP2D6 (amph, 4-OH-meth, norephedrine)	CYP2D6 (MDA, DHMA) CYP1A2 (MDA, DHMA) CYP3A4 (DHMA)	Plasma and liver esterases (EME) CYP3A4 (norcocaine)	Carboxylesterase 1 (Mor)	UGT 2B7 (M3G, M6G); UGT1A1, 1A3, 1A6, 1A9, 1A10 (M3G only) CYP3A4 and CYP2C8 (normorphine)	CYP2C9 and CYP2C19, (11-OH-THC) CYP3A4 (8-β-OH-THC)
Main metabolite or biomarker	Acetaldehyde, Ethyl glucuronide	Amph	Amph (10%)	DHMA and MDA	BE; EME	Mor; 6-MAM; M6G	M6G	11-nor-9-carboxy-THC
Active metabolites	−	4-OH-Amph 4-OH-norephedrine	Amph 4-OH-Meth	MDA	Norcocaine, Cocaethylene	Mor, 6-MAM, M6G	M6G	11-OH-THC, 8-β-OH THC
Detection time in urine	24−96 h	1−4 days	3−5 days	1−5 days	BE 2−4 days. Up to 10 days (chronic)	2−3 days	2−3 days	2−3 days 4−5 weeks (chronic)
LD50 ($mg\,kg^{-1}$)	Rats (oral) 5628−10300 Mice (oral) 3450	Rats (sc) 180 Mice (oral) 24.2 Rats (oral) 55	Rats (ip) 70	Mice (ip) 97 Rats (ip) 49 Guinea pigs (ip) 98	Rats (iv) 17.5	Mice (iv) 21.8	Mice (iv) 226−318	Rats (oral) 730−1270 (iv) 40 (inhalation) 105.7

Note: For details see text.

from the first to second centuries AD. The popularity of cannabis recreational use by young people on both sides of the Atlantic was closely linked to the protest and rebellion associated with the 1960s generation. Cannabis contains more than 400 different chemical compounds, but the main psychoactive chemical compound is the delta-9-tetrahydrocannabinol (THC). Cannabis is used recreationally as a psychoactive drug under unprocessed or processed forms. In the streets, cannabis may be found in the form of leaves or small stems (known as marijuana, bhang, dagga, or kif), female flower heads (sensimilla), as resin (known as hashish, hash, charas, or polm), or oil (alcoholic resin extract). These forms have different levels of purity, ranging from 1 to 60% THC. In recent years, there has been a large increase in the consumption of home-grown cannabis – often using modern strains of plants yielding a high THC content. Cannabis is consumed in many different ways, most of which involve inhaling vaporized cannabinoids (smoke) from small pipes, paper-wrapped joints, or tobacco leaf-wrapped blunts. Cannabis may also be ingested orally in foods and drinks.

Clinically, cannabis may be used in the treatment of anorexia associated with weight loss in patients with AIDS, and to treat mild to moderate nausea and vomiting associated with cancer chemotherapy. Recreationally, marijuana is used for its mood altering effects, euphoria, and relaxation.

Routes of Exposure

The most common way of consuming marijuana and hashish is through inhalation. The inhaled smoke of one cigarette (joint) may contain 0.5–0.7 g of delta-9-THC, but a common dose generally contains 5–25 mg THC. Marijuana can be smoked directly or through small pipes or 'bongs.' The oral route is the usual route of administration for medical purposes. Cannabis may also be ingested orally in foods and drinks for recreational purposes.

Pharmacokinetics

- Absorption

Absorption by the oral route of administration is slow, with low, delayed peak THC levels. Bioavailability is reduced following oral ingestion (4–12%) due to extensive first pass metabolism.

Smoking marijuana results in rapid absorption, with a bioavailability of 8–50%. Peak THC plasma concentrations generally occur during the act of smoking. Typical peak plasma concentrations range from 100 to 200 ng ml^{-1} and drop below 5 ng ml^{-1} less than 3 h after smoking. The minimum plasma concentration of THC, which produces psychotropic effects, was reported as 25 ng ml^{-1}.

- Biodistribution

THC is highly lipophilic and thus widely distributed throughout the organism, with high concentrations accumulating in fatty tissues that are then slowly released into the circulation. The volume of distribution is about 4–14 l kg^{-1}.

- Biological half-life

The half-life of THC is about 3 days. Plasma concentrations of THC and its metabolite 11-hydroxy-THC decline in a few minutes, due to their redistribution to fatty tissues. After that the decline is slow, with a half-life of 30 h. The half-life may be increased in chronic users, from 2.9 to 5.0 days.

- Metabolism

THC is primarily metabolized by CYP2C9 and CYP2C19 to 11-hydroxy-THC, which has equipotent psychoactivity. The 11-hydroxy-THC is then rapidly metabolized to the 11-nor-9-carboxy-THC (THC-COOH), which is not psychoactive, and then to non-cannabinoid metabolites such as terpenes and alkenes. CYP3A4 in human liver catalyzes the oxidation of the 7- or 8-position of THC.

THC and its metabolites persist in human plasma for several days or weeks. Chronic marijuana smokers metabolize THC more rapidly than nonsmokers.

- Excretion

THC is rapidly and extensively metabolized with very little THC being excreted unchanged from the body. A majority of THC is excreted in the feces (~65%), with approximately 30% of the THC being eliminated in the urine, as conjugated glucuronic acids and free THC hydroxylated metabolites.

Metabolites can be detected in urine even 2–3 days after one exposure and, in cases of chronic use, after 4–5 weeks of abstinence.

Pharmacology and Toxicology

- Pharmacodynamics (mode of action)

THC binds to the cannabinoid receptors CB1 and CB2, and interferes with important endogenous cannabinoid neurotransmitter systems. CB1 exists mainly in the brain and in the nerve terminals innervating the gastrointestinal system, whereas CB2 is expressed mostly in immune cells.

Cannabis affects the CNS by activating CB1 receptors located on excitatory and inhibitory nerve terminals. Receptor distribution correlates with brain areas involved in physiological, psychomotor, and cognitive effects.

Repetitive use of cannabis is explained by the interaction of THC with presynaptic CB1 receptors located at

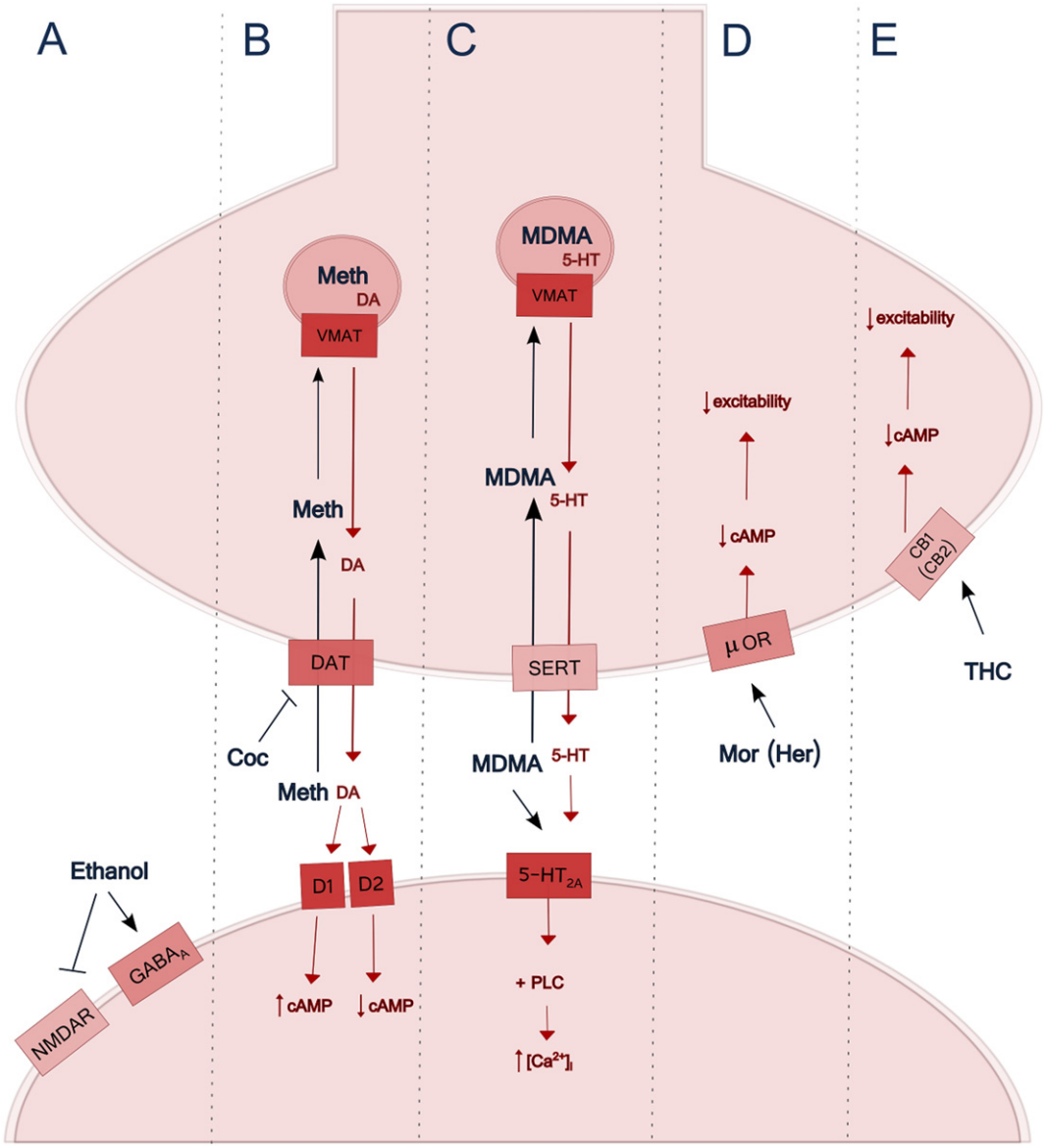


FIGURE 17.2 Molecular targets and main consequences of the actions of drugs of abuse in nerve terminals. **A)** Ethanol shares structural similarities with the neurotransmitters GABA and glutamate and interferes with $GABA_A$ and NMDA receptors, depressing neuronal activity. Methamphetamine (Meth), cocaine (Coc), and MDMA interact directly with monoaminergic nerve terminals. **B)** Meth is a substrate for the dopamine transporter (DAT), entering into the nerve terminal. There, it impairs the storage of dopamine (DA) in synaptic vesicles through VMAT, increasing DA cytosolic concentration and inducing the DAT reverse transport. Coc inhibits DA reuptake by the DAT. Both Meth and Coc lead to an increase in synaptic DA via direct effects in the DAergic nerve terminal, leading to increased activation of DA receptors. **C)** MDMA is a substrate for the serotonin transporter (SERT), being carried into serotonin (5-HT) terminals and leading to impairment of 5-HT storage in synaptic vesicles: This results in the release of 5-HT to the cytosol and reverse transport by SERT, thus increasing synaptic levels of 5-HT. MDMA is also an agonist of 5-HT2A receptors. **D)** Heroin is metabolized into morphine, which is a preferential agonist of μ-opioid receptors. **E)** Delta-9-tetrahydrocannabinol (THC), the psychoactive component of marijuana, activates cannabinoid receptors CB1 and CB2. Activation of μ-opioid and/or cannabinoid receptors leads to a decrease in adenylate cyclase activity, leading to decreases in electrical excitability and neurotransmitter release.

inhibitory GABAergic interneurons in the reward pathway, leading to decreased gamma-aminobutyric acid (GABA) release. This causes disinhibition of dopaminergic neurons and leads to an increase in synaptic dopamine concentrations, similarly to what happens with opioid drugs of abuse.

- Toxicity

THC produces alterations in motor behavior, perception, cognition, memory, learning, endocrine function, food intake, and regulation of body temperature. Long-term use of cannabis has been associated with the occurrence of psychotic episodes.

Activation of CB1 receptors causes profound coronary and cerebral vasodilation and hypotension. An increase in heart rate, usually accompanied by a mild increase in systolic pressure, is generally observed after THC use.

Lethal blood concentration of THC is within 0.180–0.315 mg l^{-1}, and the lethal dose is over 15 g.

In rats, LD50 (oral) of cannabis is within 730–1270 mg kg^{-1}, LD50 (iv) is about 40 mg kg^{-1}, and LD50 (inhalation) is 105.7 mg kg^{-1}. The LD50 in mice is 22 mg kg^{-1} (oral) and in monkeys is 130 mg kg^{-1} (iv).

CONCLUSIONS

Drugs of abuse, such as alcohol, amphetamines, cocaine, heroin, and cannabis have distinct toxicological properties and cause severe medical complications. Table 17.2 summarizes and compares the main chemical, pharmacological, and toxicological properties of these drugs, which share chemical similarities with neuronal active endogenous compounds and are thus psychoactive. All drugs of abuse affect neurotransmission, by interfering with neurotransmitter receptors and transporters. The molecular targets of these drugs of abuse in nerve terminals are represented in Fig. 17.2.

SEE ALSO

Alcohol Use Disorders, Heroin Addiction, Cocaine Addiction, Marijuana Use and Abuse, Methamphetamine Addiction, Hallucinogens, Ecstasy/MDMA

List of Abbreviations

ADHD	attention deficit hyperactivity disorder
BE	benzoylecgonine
CNS	central nervous system
CYP	cytochrome P450
DAT	dopamine transporter
DHA	3,4-dihydroxyamphetamine
DHMA	3,4-dihydroxymethamphetamine
DOB	2,5-dimethoxy-4-bromoamphetamine
EME	ecgonine methyl ester
FAS	fetal alcohol syndrome
GABA	gamma-minobutyric acid
5-HT	serotonin
LD50	median lethal dose
LSD	lysergic acid
M3G	morphine-3-glucuronide
M6G	morphine-6-glucuronide
6-MAM	6-monoacetyl morphine
MDA	3,4-methylenedioxyamphetamine
MDMA	methylenedioxymethamphetamine
4-MTA	4-methylthioamphetamine
SERT	serotonin transporter
THC	9-tetrahydrocannabinol
VMAT-2	vesicular monoamine transporter-2

Glossary

Bioavailability the proportion of drug absorbed into the systemic circulation.

Biodistribution the extent of distribution of a drug throughout the body.

Biotransformation chemical modification(s) made by an organism on a chemical compound.

First pass effect the loss of drug, following oral administration, due to hepatic metabolism, before it reaches systemic circulation.

Hyponatremia an electrolyte disturbance in which the sodium concentration in the plasma is lower than normal.

Median lethal dose (LD50) dose of a drug that kills half (50%) of the population tested (LD = lethal dose).

Plasma half-life a measure of the elimination rate, indicating the time it takes for the plasma concentration of a drug to reach half of its original concentration.

Parenteral route of administration independent of the gastrointestinal system.

Pharmacodynamics biochemical and physiological effects of drugs on the organism, including the mechanisms of drug action and the structure–activity relationship.

Pharmacokinetics study of the fate of substances administered externally to a living organism.

Rhabdomyolysis the destruction or degeneration of skeletal muscle tissue that is accompanied by the release of muscle cell contents (such as myoglobin and potassium) into the bloodstream.

Volume of distribution a measure of the volume in which the total amount of drug used would need to be uniformly distributed to produce the observed blood concentration.

Xenobiotic a chemical compound that is foreign to a living organism.

Further Reading

Brunton, L., Blumenthal, D., Buxton, I., Parker, K., 2007. The Goodman and Gilman's Manual of Pharmacology and Therapeutics. In: Brunton, L., Blumenthal, D., Buxton, I., Parker, K. (Eds.), McGraw-Hill Professional, first ed. New York, NY, USA.

Capela, J.P., Carmo, H., Remião, F., et al., 2009. Molecular and cellular mechanisms of ecstasy-induced neurotoxicity: an overview. Molecular Neurobiology 39 (3), 210–271.

Carvalho, M., Carmo, H., Costa, V.M., et al., 2012. Toxicity of amphetamines: an update. Archives of Toxicology 2012 Mar 6. [Epub ahead of print]; DOI:10.1007/s00204-012-0815-5.

Couper, F.J., Logan, B.K., Drugs and Human Performance Fact Sheets (http://www.nhtsa.gov/people/injury/research/job185drugs/technical-page.htm).

Cunha-Oliveira, T., Rego, A.C., Oliveira, C.R., 2008. Cellular and molecular mechanisms involved in the neurotoxicity of opioid and psychostimulant drugs. Brain Research Reviews 58 (1), 192–208.

Gable, R.S., 2004. Comparison of acute lethal toxicity of commonly abused psychoactive substances. Addiction 99, 686–696.

Iversen, L., 2008. The Science of Marijuana, second ed. Oxford University Press, Inc., New York, NY, USA.

Karch, S.B., 2007. In: Karch, Steven B. (Ed.), Drug Abuse Handbook, second ed. CRC Press, Boca Raton, FL, USA.

Karch, S.B., 2009. In: Karch, Steven B. (Ed.), Pathology of Drug Abuse, fourth ed. CRC Press, Boca Raton, FL, USA.

Qiunn, D.I., Wodak, A., Day, R.O., 1997. Pharmacokinetic and pharmacodynamics principles of illicit drug use and treatment of illicit drug users. Clinical Pharmacokinetics 33, 344–400.

Soine, W.H., 1986. Clandestine drug synthesis. Medicinal Research Reviews 6 (41), 74.
Timbrell, J.A., 2009. In: Timbrell, John A. (Ed.), Principles of Biochemical Toxicology, fourth ed. Informa Healthcare USA, Inc., New York, NY, USA.
United Nations Office on Drug and Crime, 2010. World Drug Report. http://www.unodc.org/documents/wdr/WDR_2010/World_Drug_Report_2010_lo-res.pdf.
World Health Organization, 2004. Neuroscience of Psychoactive Substance Use and Dependence, Geneva, Switzerland.
World Health Organization, 2011. Global Status Report on Alcohol and Health. http://www.who.int/entity/substance_abuse/publications/global_alcohol_report/msbgsruprofiles.pdf.

Relevant Websites

http://www.emcdda.europa.eu/publications/drug-profiles – Drug profiles (European Monitoring Center for Drugs and Drug Addiction).
http://www.ecstasydata.org/ – Ecstasy Data
http://www.niaaa.nih.gov/Publications/Pages/default.aspx – National institute on alcohol abuse and alcoholism (NIAAA).
http://www.nida.nih.gov/drugpages/ – National institute on drug abuse (NIDA) drug pages.
http://www.inchem.org/pages/pims.html – Poisons Information Monographs (International Programme on Chemical Safety).
http://toxnet.nlm.nih.gov/ – Toxnet.

CHAPTER

18

Tolerance and Withdrawal

Baine B. Craft, M. Kathleen B. Lustyk
Seattle Pacific University, Seattle, WA, USA

OUTLINE

TOLERANCE

As a construct, tolerance can be defined as a decrease in the magnitude of drug effects over the course of successive administrations of the same dose of drug. Stated another way, when a larger dose of a drug is needed to achieve a desired effect, tolerance has occurred. From a clinical perspective, tolerance is one of the seven potential symptoms or criteria used to diagnose substance dependence; however, dependence is not necessary for tolerance to occur.

Drug tolerance can be operationally defined as a shift in the dose–response curve to the right. As shown in Fig. 18.1, when the magnitude of the drug effect (*y*-axis) is plotted against the dose of the drug (*x*-axis), the dose response of a given drug follows a similar curvilinear trajectory irrespective of the dose. However, the dose at which a person achieves the desired drug effects is much less with first time use than the dose of drug needed with subsequent use. The difference in the dose needed to reach the desired magnitude of drug effect is explained by tolerance. This operational definition allows for tolerance to be mathematically quantified and empirically studied.

Tolerance can be defined more specifically as acute and chronic tolerance. Acute tolerance is the result of a single exposure to a substance whereby the drug effect decreases over time. Procedurally, the development and effects of acute tolerance have been studied when a single dose of a drug is administered or when a single dose of a drug is administered over a prolonged period of time during a single session. In ethanol studies, for example, acute tolerance occurs when the effects of ethanol, which cannot be explained by the half-life of the drug, decrease over time and exposure to ethanol during a single session. As a result of exposure to ethanol, more impairment in behaviors and cognitive performance occurs during the same ascending blood ethanol concentrations. However, once compensatory reactions that result in tolerance have occurred, fewer errors in behavioral and cognitive performance occur during descending blood ethanol concentrations.

Where acute tolerance is the result of a single session or administration of a drug over a brief amount of time (e.g. minutes or hours), chronic tolerance is due to several sessions of exposure to a drug across a longer amount of time such as days. Chronic tolerance more closely mimics the development of tolerance that is common in human drug dependence. For example, human substance abuse follows a typical pattern of relatively brief but repeated drug administration over days or weeks. Evidence from some studies suggest a relationship between acute and chronic tolerance such that associations with a particular drug that occur during acute tolerance facilitate the tolerance that occurs with chronic administration of a drug.

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00018-8

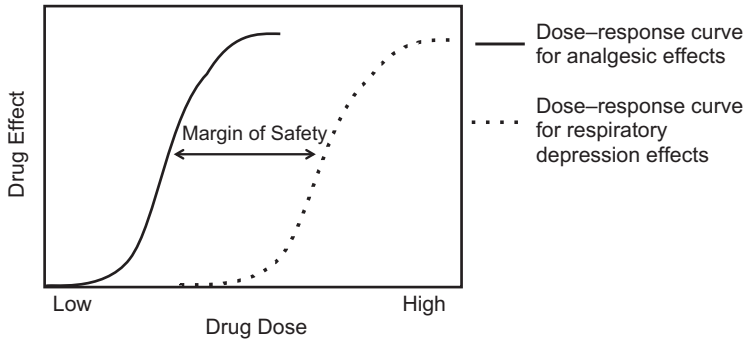


FIGURE 18.1 Margin of safety between therapeutic and harmful doses of opiates.

Research studies suggest that no single mechanism predicts tolerance. Rather, tolerance is due to various combinations of mechanisms, which may or may not be dependent upon the type of drug used. These mechanisms can be broadly categorized into biological and behavioral mechanisms or forms of tolerance. For example, research has shown that tolerance can result from changes in the metabolism of a drug, a decrease or increase in the responsiveness of neurotransmitter receptors, or as the result of learning. In addition, changes in the dose, type of drug consumed, or individual (i.e. genetic) factors can result in differences in the development of tolerance.

Biological Mechanisms

Biologically, tolerance is the result of adaptations or responses to changes in homeostasis. Homeostasis is a term coined by Walter Cannon in the early 1900s to mean internal stability despite environmental change. According to Cannon, environmental change could be either endogenous (i.e. within the individual) or exogenous (i.e. outside the individual). When confronted with the presence of an exogenous or endogenous chemical substance, the body responds in different ways to return to homeostasis. Moreover, both endogenous and exogenous factors can contribute to the homeostatic adaptations that result in tolerance. For example, a primary drug effect of alcohol use is hypothermia (i.e. a decrease in body temperature). Due to the hypothermic effect, a compensatory reaction occurs such that the body compensates for the presence of alcohol in the body by increasing body temperature through thermoregulatory mechanisms. While the compensation for changes in homeostasis is adaptive, it also underlies tolerance. In other words, tolerance is caused by compensatory reactions that oppose the primary drug effects.

Types of compensatory reactions include pharmacodynamics or metabolic adaptations, where pharmacodynamics refers to changes in the body in response to a drug. An example is when the tissue adapts to the drug. This adaptation can occur at the molecular level in the form of receptor binding site modification. If, for example, the effect of a drug is to increase the actions of a certain neurochemical, the body may respond by decreasing the number of active ligand-binding sites on the receptors for that neurochemical. More specifically, using positron emission tomography (PET), Michael Nader and his colleagues at Wake Forest University Medical School showed a reduction in dopamine D_2 receptor availability in response to cocaine abuse.

More complex adaptations at the molecular level can also occur such as those involving second messenger systems, which ultimately diminish receptor–ligand binding efficiency and/or the effects of receptor–ligand binding on cellular activity. For example, research with knockout transgenic mice lacking the opiate μ receptor shows that these animals are immune to the rewarding actions of morphine. Given that μ receptor actions are mediated by G-protein second messengers with various effects, myriad functional changes are possible due to tolerance at these sites. Resultantly, these biological responses compensate for the increased neurochemical activity brought on by the drug.

Given the properties of neural receptors, a drug-induced adaptation at the receptor level may result in cross tolerance or tolerance specificity. Cross tolerance occurs when one drug produces tolerance to another drug that acts via the same biological mechanism. For example, alcohol-dependent individuals may experience cross tolerance with benzodiazepines such that alcoholics may fail to obtain the anxiolytic effects from benzodiazepines. Both alcohol and benzodiazepines are central nervous system depressants with a primary site of action being the gamma-aminobutyric acid (GABA) receptor complex. By inducing changes that make ligand–receptor binding more efficacious, alcohol and benzodiazepines possess pharmacologic synergy and cross tolerance. While this may reduce the potential anxiolytic effects of benzodiazepines in the alcoholics, it does allow benzodiazepines to be used for alcohol detoxification. Because benzodiazepine dosage can be more tightly regimented than alcohol, step-down dosing

with benzodiazepines can ease the suffering experienced during rehabilitation from alcoholism.

Conversely, tolerance specificity is also possible. This occurs when specific neural adaptations to a drug result in tolerance to some drug effects but not others. For example, the primary effects of opiates on analgesia and euphoria are subject to tolerance whereas opiate effects on autonomic nervous system driven changes such as respiratory depression are not. These processes stand in stark contrast to sensitization, which can be thought of as the opposite of tolerance.

Sensitization occurs when drug effects become enhanced. Sensitization to the dopaminergic response in the nucleus accumbens has been observed in rats repeatedly exposed to amphetamines. That is to say, with repeated doses of amphetamines, the drug increases its ability to increase extracellular levels of dopamine in the pleasure circuit. Such findings have important clinical implications as they may help us understand an addict's progressive increases in motivation to seek out and use drugs. However, research attempts to demonstrate sensitization in multiple animal models and in humans are equivocal. As such, some researchers suggest that a model of drug abuse that includes processes of sensitization is of limited value. Interested readers are referred to the review of the topic by Paul Venzia cited in the reference section of this chapter.

With tolerance, metabolic adaptations can also occur. An example is drug dispositional tolerance. This can occur when the liver develops a more efficient clearance rate for the drug. For example, with repeated use of barbiturates, the concentration of microsomal liver enzymes responsible for metabolism of the drug increases. As a result, the drug can be more rapidly eliminated from the body. Other metabolic changes may involve drug absorption, distribution, and/or excretion, which can also result in dispositional tolerance.

Together, pharmacodynamics and drug dispositional tolerance counteract the desired effects of the drug resulting in the need for more of the drug to achieve the desired effect or high.

Behavioral Mechanisms

Using standard Pavlovian conditioning procedures, Dr Shepard Siegel and other researchers have demonstrated the powerful role of Pavlovian conditioning in the development of tolerance. Tolerance that results from learning or behavioral tolerance is a decrease in drug efficacy that is due to behavioral or associative mechanisms. This type of tolerance has been referred to as situation-specific or environment-dependent tolerance due to the influence of the situational or environmental cues on compensatory responses.

In situation-specific tolerance, the presence of a drug in the body is defined as an unconditioned stimulus, and the primary drug effect is the unconditioned response. After repeated pairings of the drug with neutral stimuli in a specific situation or environment (e.g. drug paraphernalia), classical conditioning occurs such that the previously neutral stimuli within the environment become conditioned stimuli that elicit a conditioned response. Conditioned stimuli (e.g. drug paraphernalia) that have been associated with a particular drug, when presented in the absence of the drug, elicit compensatory reactions; these reactions are called conditioned compensatory reactions (see Table 18.2).

Conditioned compensatory reactions serve to counteract primary drug effects. For example, using a rat model, if subjects were given alcohol in the presence of cues previously associated with alcohol (i.e. stimuli in the room), Siegel observed a conditioned compensatory reaction such that the cues caused an increase in body temperature. By increasing body temperature, the stimuli that had been paired with alcohol elicited a compensatory reaction to counteract the hypothermic effects of alcohol. However, if rats were exposed to alcohol in the absence of cues previously associated with alcohol (i.e. in a different room where they had not previously been given alcohol), a conditioned compensatory response did not occur, and subjects' body temperature did not increase. In the previous example, cues in the animals' environment elicited a conditioned compensatory reaction. These cues can be categorized as exteroceptive cues.

External or exteroceptive as well as internal or interoceptive are two types of cues that have been shown to result in conditioned compensatory reactions. Exteroceptive cues (e.g. visual, auditory, or olfactory) are observable and external to the organism, and when paired with a drug effect, these cues will result in a conditioned compensatory reaction. For example, in the rat model of heroin addiction, Siegel demonstrated that when rats experienced the administration of a drug in a distinctive environment, subjects' behaviors were less lethargic or more tolerant to morphine's sedative effects on behavior. However, when morphine was administered in novel or different environment than previously paired with morphine, subjects were less tolerant to the effects of the drug.

Interoceptive cues are internal or unobservable, but when paired with a drug effect (similar to exteroceptive cues), these cues will result in a conditioned compensatory reaction. For example, one type of interoceptive cues are drug-onset cues. Research has shown that when associated with a particular drug effect, the initial onset or effect of a drug can cause conditioned compensatory reactions. Animal studies of interoceptive cues resulting from pairings with ethanol have shown that

subjects with prior exposure to ethanol will display conditioned compensatory reactions when only a small dose of ethanol is administered. That is, increases in rat's body temperature were observed when a small dose of ethanol was administered to subjects with prior experience with ethanol.

A second example of interoceptive cues, self-administration cues, occur when individuals are preparing to administer a drug. Whether heroin or smoking a cigarette, individuals develop patterns of behavior that precede the administration of a drug. These behavioral patterns and the presence of drug paraphernalia before and during drug administration can serve as strong conditioned stimuli that elicit conditioned compensatory reactions. For example, the process of preparing a syringe or tying a tourniquet before using heroin or the process of packing or opening a pack of cigarettes can become conditioned stimuli that elicit conditioned compensatory reactions. While the cues in the aforementioned example may seem external, they are, however, internal as the behaviors being performed are initiated by the subject. These processes will be further discussed below as we focus on withdrawal.

Tolerance and Addiction

Due to the increased dose required to achieve a particular drug effect once tolerance has occurred, tolerance plays an integral role in the development of an addiction. These effects are particularly dependent on alterations in the neurocircuitry that allows for the pleasurable experiences of drug use. This circuitry involves dopaminergic neurons originating from a region in the lower brain called the ventral tegmental area (VTA). These neurons give rise to mesolimbic and mesocortical projections implicated in reward. Commonly abused substances such as nicotine, alcohol, and opiates affect the release of dopamine and/or the levels of dopamine within the complex circuitry known as the mesocorticolimbic dopaminergic system. Meso refers to the mesencephalon or midbrain, which as its name implies lies toward the middle of the brain. The mesencephalon is divided into the tectum, or roof, and the tegmentum, or floor. Within the ventral part of the tegmentum lies the VTA. One target of the dopamine fibers originating in the VTA is a limbic system structure referred to as the nucleus accumbens. The nucleus accumbens is a region located further up in the brain in an area called the ventral striatum.

Much research has shown that all drugs of abuse increase the release of dopamine in the nucleus accumbens. From animal studies, we learn that the increased release of dopamine is pleasurable as animals will exhibit high rates of intracranial self-stimulation to cause fibers in this region to release dopamine. That is, after having their brains implanted with an electrode that will provide a stimulating electric current when the animal presses a lever, animals will press the lever at extremely high rates; even to the point of physical exhaustion whereby they can no longer press the lever. Moreover, rats will choose to press the lever to receive the stimulation to their brains over eating palatable foods or engaging in copulatory behavior. It is also the case that the administration of dopamine receptor blockers to synaptic regions of this circuit reduce intracranial self-stimulation in the rat model. Thus, irrespective of the primary effects of an abused substance such as alcohol with its central nervous system depressant effect, substance use is rewarding or pleasurable because it increases dopamine within the mesocorticolimbic circuit. Because of the saliency of the pleasure experienced with activation of this circuit, researchers have termed this the pleasure circuit.

Advancements in neuroimaging have allowed for the investigation of the pleasure circuit in humans. For example, PET has been used to show increased dopamine release in the ventral striatum in response to alcohol and amphetamines. One property of the pleasure circuit neurobiology that is detrimental to the substance user is system down-regulation. Simply put, with repeated exposure to a substance, the pleasure circuit is less activated by the substance. At the cellular level, this down-regulation reflects tolerance, which is subjectively experienced as the failure of the substance to meet the expectations desired. To meet expectations, additional use is needed, which in turn results in more cellular adaptation and system down-regulation. At this point, the abuser has become dependent on the substance. For readers interested in more details on the cellular response to drug dependence, we recommend the 2007 review by Kauer and Malenka.

Amidst the receptor adaptation taking place within the nucleus accumbens, another pathway from the VTA to the dorsal striatum, which includes the caudate and putamen, comes into play further driving repeated substance use. This pathway, again using dopamine as the chemical messenger, is implicated in a simple form of instrumental or reward-based learning. This kind of learning is behind the behaviors we call habits. In one sense, drug addiction is habitual or compulsive use resultant from a loss of control over voluntary use. Research in monkeys supports this model of progression at the neural level with dopaminergic activation beginning in the nucleus accumbens (ventral striatum) with early cocaine administration and progressing upward to activate the dorsal striatum (caudate and putamen). This latter circuit has been termed the habit circuit.

The research of Rajita Sinha and her colleagues at Yale University reveals that in humans, this habit circuit is activated during stress. Thus, for a person with

a substance use disorder, stress becomes a trigger for a habitual response which in this case is substance use. The motivation for the habit or use is entrained at the cellular level. Given that the activator of this circuit is stress, motivation to use may be compounded by craving in those with substance use disorders. This is because craving can be a stressor in and of itself. Craving can result from withdrawal sensations between bouts of use. In response to craving, a person may habitually use to avoid the negative sensations resultant from withdrawal. Thus, the voluntary and conscious choice to ingest a drug such as enjoying a happy hour drink with colleagues following a stressful board meeting, may become a habitual response to stress over time. We discuss the role of stress and withdrawal in more detail later in this chapter.

Fortunately, inhibitory pathways from the prefrontal cortex (PFC) can modulate the function of the habit circuit. The PFC consists of many brain regions that are located behind the forehead. Inhibitory pathways from the PFC can be recruited to regulate habits. Many people possess what they call annoying habits. A good example is nail biting. Some people who bite their nails find that they often engage in the habitual behavior mindlessly. Operating on "automatic pilot," they can bite the nail down lower than the nail bed which can result in later discomfort. However, when they bring their conscious attention to their behavior, they find that they can temper the motivation to bite their nails. This is often true of other habitual behaviors as well. To refrain in this manner, one uses their PFC. The PFC is often called the brain "executive." This is because of its role in executive functions which are regulatory in nature. Some examples are emotion regulation, cognitive flexibility, planning, and impulse control. The latter is particularly relevant to the development of addiction. Using the neurotransmitter GABA, fibers from the PFC can inhibit neural activity in various subcortical regions including the dorsal striatum, a region within the habit circuit. Unfortunately, substance abuse has been shown to result in the loss of PFC gray matter and the development of functional hypofrontality. That is to say, the frontal regions of the PFC lose their ability to function fully. It follows then that a loss of PFC function would result in a loss of inhibition over the habit circuit and mindless habitual behavior be it nail biting or substance use. Moreover, the loss of PFC function could contribute to what addictions expert, G. Alan Marlatt, PhD, called the the problem of immediate gratification (PIG). When an addict has an urge or suffers craving, they may lack the executive function of impulse control thus succumbing to the urges and cravings and eventually relapse.

G. Alan Marlatt wrote extensively on the habitual nature of drug use and the powerful role of cue association in drug craving and potential drug relapse. In developing what is now a part of the canon for all clinical professionals, Relapse Prevention Therapy teaches people the skills aimed at dealing with triggers and high-risk situations that have a strong potential to produce urges and cravings that may lead to relapse. One way of dealing with PFC deficits is to tap metacognitive processes that introduce a kind of metal space between the cue or trigger and the urge or cravings. One metacognitive technique developed by Alan Marlatt and his colleagues Drs Sarah Bowen and Neha Chawla of the University of Washington Addictive Behaviors Research Center is the SOBER breathing space (SOBER breathing space is one of several mindfulness-based techniques described in detail in Mindfulness-Based Relapse Prevention for Addictive Behaviors: A Clinician's Guide by Sarah Bowen, Neha Chawla, and G. Alan Marlatt.). This technique teaches people how to mentally take a step back from their experience and engage in thought processes that explore their experience. True to the meaning of the term metacognition, the SOBER breathing space causes one to think about their thinking. SOBER is an acronym with the S meaning stop. In other words, stop right where you are and observe (the O) how you're feeling. You can ask yourself "What is this sensation?" You can give it a name "This is craving." Recognizing that all are too familiar sensations or feelings, you mentally shift your focus to your breathing, which is the B in SOBER. The idea is to follow your breath mentally. Breathing in, breathing out, you let your breath become the center of your attention. Next, you expand (the E) your awareness and gather a better of sense of the consequences associated with the habitual act. You can ask yourself "How would it feel to take that drink?" for example. Finally, you respond (the R in SOBER) mindfully and in the present moment returning to the conscious way you once engaged the behavior before it became a habit. Thus, this metacognitive strategy brings a mindfulness component to the classic cognitive-behavioral treatment strategy – Relapse Prevention.

In addition to what neurobiological and pharmacological science teaches about the etiology and effects of tolerance, behavioral economists have described tolerance as affecting the consumption of drugs relative to engaging in alternative behaviors. From this perspective, the science of behavior economics borrows from traditional economic theory and applies it to behavioral processes such that behaviors are viewed as having a relative utility or benefit. The consumer (i.e. the individual performing behaviors) is seeking to maximize the utility or benefit gained from each behavior in which they engage. While every behavior delivers a return, potentially with a variable rate of gain, each behavior also has a cost. Thus, according to behavior economics,

any behavior or patterns of behavior, including substance abuse, can be best described as being the result of the interplay between the cost of performing a behavior and the relative rate of gain from performing a behavior. Note that the cost of performing a behavior and the rate of gain returned from a behavior is not only monetary but also could be physical (e.g. effort) or psychological (e.g. stress).

For example, the price of a drug may be varied such that time, energy, social and familial relationships, and money must be paid in order to consume the drug. In fact, for many addicts, the cost of abusing a drug requires, for example, the majority of their time, energy, social and familial relationships, and money such that he or she does little else than seek out and use the drug. At the same time, the return from some drugs carries a rate of gain that has been described as being unparalleled to any other experience. For example, some heroin abusers describe the euphoria of drug effects as being more rewarding than any other experience (i.e. the effects are immediate and without any vagaries regarding its efficacy). In other words, unlike other behaviors in which a person engages where the return might be uncertain or vary (such as a social gathering), when it comes to drug use, there is no question as to the effects of the drug (i.e. the drug always works and it works quickly).

According to behavior economics, as the rate of gain from a behavior decreases and as the price of a behavior increases, as with tolerance, the way an individual allocates behaviors or chooses behaviors will also change. For example, with substance abuse, the effects of tolerance result in a change in the rate of gain from drug consumption as well as the price required to achieve the desired drug effect. Specifically, as an abuser uses their drug of choice and the biological and associative effects of tolerance are established, the rate of gain from the drug decreases. To increase the gain, more drug is needed and this is costly. Thus, the decrement in the return from drug coincides with an increment in the price of the drug needed to achieve the desired effect. Given the decrease in the rate of gain and resultant increase in price, tolerance should cause the drug seeking and drug use behaviors to decrease. However, whether or not the individual will continue seeking out and using drugs at the increased price is due to a variety of factors such as further changes in the price of drug use (i.e. elasticity of demand) and the temporal discounting of alternative behaviors (for a more detailed review see Vuchinich and Heather, 2003.).

Briefly, elasticity of demand describes changes in behaviors that result from fluctuations in the price of performing a behavior. In the case of substance abuse, the consumption of a drug is inversely related to the price of consuming the drug. In other words, there is flexibility or elasticity in the demand to consume substances such that some behaviors associated with drug use are more elastic than others. Whether the demand is inelastic or elastic depends on the price of the drug. When the price changes, as it does with tolerance, changes in the demand for a drug use can determine drug use. For example, research has shown that if the price of illicit drugs, cigarettes, or alcohol is low, the demand for the drug is inelastic. The effects of this inelastic demand on behavior is that the consumer will pay only small increments in cost (e.g. time, relationships, energy, money, etc.) in order to consume the drug. However, if the cost of the drug increases substantially such that the demand becomes elastic, the consumer won't pay the large cost but will instead choose to perform behaviors that are substitutes for the drug (e.g. seeking a drug or behavior that will serve as a substitute for the preferred substance) or alternatives to using the drug such as a nondrug alternative. A nondrug alternative can be defined as any behavior that could potentially serve as an alternative for drug use. For example, social, marital, and familial support has been shown to serve as alternatives to drug use such that if social support is available, a decrease in drug use will occur.

In most substance abuse situations, the increase in price that results from tolerance does not result in a decrease in consumption but rather the increase in price contributes to an increased consumption of the drug at a higher price. The elasticity of demand for drugs, despite increased tolerance, can be described as being the result of a decrease in value associated with nondrug alternatives due to temporal differences associated with drug use and nondrug use behaviors. Despite the value associated with nondrug alternatives such as spending time with family, the potential delay of gratification associated with such alternative behaviors stands in stark contrast to the immediate gratification of a drug. Moreover, as mentioned before, there are none of the vagaries about the drug's effectiveness; it always works. However, social and/or family gathering are not a guarantee for pleasure. Clinically, this is what the late addictions expert G. Alan Marlatt wrote of when he referred to the PIG. This choice for an immediate but smaller reward over a larger but delayed alternative can be described as being the result of temporal discounting.

In situations where temporal discounting occurs, the immediate return of using a substance outweighs the return from nondrug alternatives due to delay. For example, if two alternatives (i.e. a drug and nondrug alternatives) are presented simultaneously with no delay associated with the alternatives, an individual will choose the option that yields the highest rate of gain. However, when the delay to a reward is longer, as is typical of nondrug alternatives, the value

associated with that reward decreases or is discounted. Due to the temporal discounting of nondrug alternatives, individuals choose to engage in drug use because of the immediate reward associated with the drug, despite the increases in price that result from the development of tolerance to the drug.

WITHDRAWAL

The adaptive biological changes that underlie drug tolerance also underlie drug withdrawal. Withdrawal is a general term used to describe a cluster of symptoms that result from adverse biological reactions to the lack of drug in the system (Table 18.1). With tolerance, the system becomes habituated to having the drug present. When the drug is removed from the system, withdrawal effects (i.e. the adverse reactions) occur. Withdrawal effects are often the opposite of the drug effects. Thus, withdrawal symptoms are drug specific. For example, a person with a seizure disorder taking an anticonvulsant such as Phenytoin, is likely to experience seizure activity if the drug is stopped suddenly. This is because the goal of treatment with Phenytoin is to develop tolerance or physical adaptation so that a certain level of the drug is present in the system at all times. Thus, when the Phenytoin level drops, withdrawal occurs. Withdrawal effects in this instance are the opposite of the drug effects. As such, anticonvulsant effects give way to seizure.

With drugs of abuse however, withdrawal also results in a decrease of dopamine in the nucleus accumbens demonstrating a lack of drug specificity on withdrawal. The loss of dopamine within the pleasure circuit has negative consequences on one's ability to experience pleasure and as such may drive behavior to replenish dopamine through use. This is particularly likely during the initial or acute withdrawal phase, when symptoms are intense. Yet symptom intensity is not ubiquitous but rather depends upon the type of drug abused, the duration of abuse, and whether the person quits the drug suddenly or whether he or she steps down off the drug in a systematic fashion that allows for system titration. Still, the presence of withdrawal symptoms following cessation of the drug indicates that physical dependency to the drug has occurred.

A classic example of drug withdrawal is opiate withdrawal. Opiates exert varied effects due to their actions throughout the central nervous system. The primary therapeutic effects of opiates are analgesia and sedation. However, when rapid entry into the body occurs at a dose higher than therapeutic range, a profound and rapid high/euphoria is experienced. This is often referred to as the "rush." Through it's actions on the pain inhibitory pathway originating in the periaqueductal gray matter of the midbrain, opiates exert their analgesic effects. Actions on the preoptic area of the hypothalamus yield a warm flushing of the skin. Sedation is experienced via opiates actions on the mesencephalic reticular formation. The drowsy and clouded mental function are often referred to by users as "on the nod" and is accompanied by a feeling of heavy extremities and muscle relaxation. Also, opiates exert a whole host of effects on the autonomic nervous system (ANS) that can pose life threatening withdrawal symptoms when a person who is tolerant to opiates suddenly stops taking the drug. The primary effects of opiates on the ANS include slowed respiration, slowed heart rate, as well as inhibitory effects on smooth muscle including that which makes up the gastrointestinal (GI) system.

With abuse, each of these effects results in a compensatory reaction opposite to the primary effect of the drug (see Table 18.2). Thus, analgesia is countered by pain,

TABLE 18.1 Symptoms of Withdrawal

Psychological	Physical
Anxiety	Sweating
Irritability	Irregular heart beat
Restlessness	Muscle tension
Insomnia	Shortness of breath
Poor concentration	Headaches
Depression	Tremor
Emotional outbursts	GI upset including nausea, vomiting, or diarrhea
Social isolation	In severe cases: seizures, heart attack, stroke, hallucinations, delirium tremens (DTs)

TABLE 18.2 Primary Effects of Heroin and Associated Compensatory Reactions

Primary Effects	Compensatory Reaction
Analgesia	Pain
Euphoria	Depression Anxiety
Muscle relaxation	Muscle tension Muscle spasms
Sedation	Insomnia
Respiratory depression	Hyperventilation
Bradycardia	Tachycardia
Hypotension	Hypertension
Constipation	Diarrhea Nausea Vomiting

often experienced as aches in the muscles and bone. Sedation is countered by agitation, restlessness, and insomnia. Heart rate, respiration, and body temperature increase and GI disturbances arise such as diarrhea, cramping, nausea, and vomiting. When the addiction is severe, the withdrawal symptoms are at their worse. For example, hyperthermia can cause perspiration with subsequent skin cooling and goose bumps. This symptom of withdrawal led to phrase "going cold turkey" as the skin appears similar to plucked poultry skin with small bumps over the surface. "Going cold turkey" has come to mean the process of suddenly stopping any addictive substance. A second example of severe withdrawal symptomatology involves painful muscle contractions and twitching such as foot kicking. This symptom of opiate withdrawal gave rise to the phrase "kicking the habit" which again has been generalized beyond its initial clinical observation to refer to the general process of getting off addictive drugs. But it's the compensatory reactions to the primary effects of opiates on the ANS that pose serious health risks to those experiencing opiate withdrawal.

Of primary concern are the compensatory reactions affecting the cardiovascular system which include increased heart rate and blood pressure. Taxing the cardiovascular system in this fashion can have dire consequences such as heart attack. Cardiac effects may be further compounded by extreme electrolyte imbalances associated with withdrawal induced vomiting and diarrhea. Such concerns render severe opiate withdrawal a medical condition worthy of care by a healthcare provider.

The compensatory reactions that manifest themselves as opiate withdrawal symptoms are subject to conditioning. That is to say, they can be learned. As was discussed earlier, the primary effects of a drug can be termed as an unconditioned stimulus using Pavlovian terminology. The unconditioned response to a drug is the set of compensatory reactions. With repeated bouts of drug use, the neutral stimuli present and the time of use and/or used to administer the drug become conditioned stimuli which can produce the conditioned response. In this case the conditioned response is the set of compensatory reactions (see Table 18.2). Keeping with the model of opiate addiction, the administration of intravenous heroin, for example, involves heating the drug, drawing the solution into a syringe, and injecting the drug. Thus there are numerous paraphernalia items needed for drug administration such as a match or lighter, a spoon or small cup to hold the drug while heating, a tourniquet to ease the process of finding a viable vein, and a syringe. Any or all of these items can become conditioned stimuli over time. Moreover, heroin users often find their use to be ritualistic. That is, they do their drug in a particular location often in secrecy as heroin use is illegal. They may douse the sound of heating the drug with music and camouflage the smell by burning incense. All of these contextual variables can become conditioned stimuli as well.

Following with the learning model then, as the addict prepares their drug and engages in their ritual, the conditioned stimuli produce the conditioned response, which is the set of compensatory reactions. This means that even before the addict administers the drug, they are in withdrawal. Thus it follows that they are feeling warm, their heart is racing, they are hyperventilating, and they feel achy. Because the compensatory mechanisms are occurring prior to drug administration, the addict will find that they can safely administer a larger dose of heroin and in fact they will need in order to experience a high. This is because the compensatory responses have increased their tolerance to the drug. But this tolerance is situational. If the addict were to use in a different environment, with different music playing, and different incense burning, the compensatory responses may not occur as intensely. If that were the case, and the user administered the same dose of drug that they usually do when they engage in their ritual in their usual context, they will likely overdose.

In the case of heroin, this overdose phenomenon is easily explained by considering the small margin of safety between the analgesic effects of the drug and the respiratory depression effects. To explain, the margin of safety is a pharmacological principal that describes the range between the therapeutic and lethal doses of a drug. As shown in Fig. 18.1, the amount of drug needed to acquire the analgesic effects of opiates (i.e. the therapeutic effects) is less than the dose that would cause respiratory depression. However, with tolerance, the dose–response curve for the analgesic effects shifts to the right, shrinking the margin of safety. Thus, the addict that is tolerant to opiates will find themselves reaching a point where a larger dose is not realistic without overdosing and killing themselves. Their search for their high, which drives their use, may never be satisfied. Instead, they use to avoid the withdrawal symptoms that occur as the drug leaves the system. Bringing the situational tolerance back into the picture, when the heroin addict engages in their drug use with their ritual, the compensatory reaction of hyperventilation makes it safe for them to take a dose of drug that may "step over" the margin of safety as depicted in the figure. The hyperventilation counters the respiratory depression effects of the drug. Now, if the addict engages in use in a novel environment without their ritual, and the compensatory hyperventilation does not occur, then the dose that steps over the margin of safety will be lethal.

When we consider that withdrawal is caused by the presence of compensatory mechanisms and that compensatory mechanisms bring about tolerance, we can see that these responses are opposite sides of the proverbial coin. On one side we have withdrawal and on the other side we have tolerance. With tolerance comes compensatory reactions which in and of themselves are withdrawal. Withdrawal occurs because of the compensatory reactions, the effects of which the addict is motivated to reduce. Thus, anything that triggers the compensatory reactions will likely drive the addict to use. One well-studied trigger is stress.

Stress and Withdrawal

A grand irony in addiction science is the fact that substances with anxiolytic effects (e.g. alcohol) actually turn on the neurochemical stress system. Specifically, drugs of abuse activate the hypothalamic–pituitary–adrenal axis (HPAA), which in turn sensitizes the pleasure circuit during early use. The research of George Koob and his colleagues at Scripps Research Institute in California points specifically to an intricate system of neural dysregulation involving both the pleasure circuit and stress circuit. In fact, research has shown that nonabstinent alcohol-dependent individuals and alcoholics undergoing acute withdrawal have high levels of cortisol, which is the end-product of the HPAA. Briefly, cortisol is a glucocorticoid released by the adrenal cortex in response to adrenocorticotropic hormone (ACTH) from the anterior pituitary located at the base of the brain. The pituitary release of ACTH is under the influence of corticotropic hormone (CRH) from the hypothalamus. So well accepted is the role of cortisol in the stress response that it is often referred to simply as the stress hormone. Once in the blood stream, cortisol promotes the release of stored glucose and works in conjunction with the sympathetic nervous system to drive the stress response often referred to as the "fight or flight" response. As it turns out, high levels of cortisol is associated with less dopaminergic response to substance use. That is, the addict gets less effect of the drug not only because of tolerance but because they are using. Thus, HPAA activation may accelerate the progression to addiction by amplifying the pleasure sensation associated with initial use and may maintain addiction by heightening the negative sensations of withdrawal associated with later use.

Much less is known about the extra-hypothalamic stress circuits. Still, research points to the amygdala with it's own CRH system as a key region involved in extra-hypothalamic stress circuitry. What George Koob and others propose is that initially the chemical milieu that results from repeated drug use promotes sensitization in the pleasure circuit. Thus, the pleasurable sensation is great as is the motivation to repeat use. As the HPAA continues to produce cortisol, the circulating levels of this stress hormone sensitize the extra-hypothalamic stress circuits, in particular the CRH system within the amygdala. Since increased CRH activity in the amygdala is associated with anxiety, this increase subsequent to substance use is proposed to underlie the anxiety-like feelings or negative affect that emerges as the substance abuse continues. The emergence of this change in affect coincides with the onset of withdrawal symptoms. Thus, what may begin as a behavior motivated to receive pleasure becomes a behavior motivated to avoid negative affect and withdrawal (for more detailed information see Koob, 2008). As such, addiction can be viewed as a complex function of tolerance plus withdrawal exponentially affected by stress.

SEE ALSO

Alcohol Use Disorders, Heroin Addiction, Overdose, Relapse and Lapse, Stress and Addiction, Metacognition in Substance Misuse

List of Abbreviations

ACTH	adrenocorticotropic hormone
ANS	autonomic nervous system
CRH	corticotropin releasing hormone
GABA	gamma-aminobutyric acid
GI	Gastrointestinal
HPAA	hypothalamic pituitary adrenal axis
PET	Positron emission tomography
PFC	prefrontal cortex
PIG	Problem of immediate gratification
SOBER	Stop, Observe, Breath, Expand, Respond exercise as part of mindfulness-based relapse prevention therapy
VTA	ventral tegmental area

Glossary

Anxiolytic antianxiety, anxiety reducing.

Conditioned compensatory reaction a classically conditioned response to a conditioned stimulus that has been associated with a drug due to repeated pairings between the conditioned stimulus and drug.

Behavioral economics describes behavior and patterns of behavior as being the result of the interplay between the cost of performing a behavior and the relative rate of gain from performing a behavior.

Drug tolerance a condition whereby a larger dose of a drug is needed to achieve a desired effect.

Drug withdrawal the appearance of symptoms opposite to those produced by a drug when drug administration ceases.

Homeostasis a state of balance or equilibrium in the body; the ability of a person or animal to maintain their internal body balance equilibrium.

Pavlovian psychopharmacology describes changes in drug efficacy that is due to behavioral or associative mechanism.

Pharmacodynamics is the study of drug effects and their mechanisms of action.

Pharmacologic synergy when two drug agents act synergistically to exert their effects on the body.

Positron emission tomography (PET) a functional brain imaging method utilizing the administration of a positron-emitting radionuclide that emits pairs of gamma rays which are detected by the scanner to construct a map of metabolic activity in the brain.

Sensitization an increase in drug effects with repeated drug administrations.

Further Reading

Belin, D., Everitt, B.J., 2008. Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral and dorsal striatum. Neuron 57, 432–441.

Blume, A.W., Marlatt, G.A., 2009. The role of executive cognitive functions in changing substance use: what we know and what we need to know. Annals of Behavioral Medicine 37, 117–125.

Boileau, I., Assaad, J.M., Pihl, R.O., Benkelfat, C., Leyton, M., Diksic, M., Tremblay, R.E., Dagher, A., 2003. Alcohol promotes dopamine release in the human nucleus accumbens. Synapse 49, 226–231.

Bowen, S., Chawla, N., Marlatt, G.A., 2010. Mindfulness-based Relapse Prevention for Addictive Behaviors: A Clinician's Guide. Guilford Press, New York, NY.

Craft, B.B., Church, A.C., Rohrbach, C.M., Bennett, J.C., 2011. The effects of reward quality on risk-senstivity in Rattus norvegicus. Behavioural Processes 88, 44–46.

Di Chiara, G., Bassareo, V., Fenu, S., De Luca, M.A., Spina, L., Cadoni, C., Acquas, E., Carboni, E., Valentini, V., Lecca, D., 2004. Dopamine and drug addiction: the nucleus accumbens shell connection. Neuropharmacology 47, 227–241.

Donovan, D.M., Marlatt, G.A. (Eds.), 2005. Assessment of Addictive Behaviors, second ed. Guilford Press, New York, NY.

Drevets, W.C., Gautier, C., Price, J.C., Kupfer, D.J., Kinahan, P.E., Grace, A.A., Price, J.L., Mathis, C.A., 2001. Amphetamine-induced dopamine release in human ventral striatum correlates with euphoria. Biological Psychiatry 49, 81–96.

Kauer, J.A., Malenka, R.C., 2007. Synaptic plasticity and addiction. Nature Reviews: Neuroscience 8, 844–858.

Koob, G.F., 2008. A role for brain stress systems in addiction. Neuron 59, 11–34.

Marlatt, G.A., Donovan, D.M. (Eds.), 2005. Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors, second ed. Guilford Press, New York, NY.

Mathalon, D.H., Pfefferbaum, A., Lim, K.O., Rosenbloom, M.J., Sullivan, E.V., 2003. Compounded brain volume deficits in schizophrenia-alcoholism comorbidity. Archives of General Psychiatry 60, 245–252.

Porrino, L.J., Lyons, D., Smith, H.R., Daunais, J.B., Nader, M.A., 2004. Cocaine self-administration produces a progressive involvement of limbic, association, and sensorimotor striatal domains. Journal of Neuroscience 24, 3554–3562.

Siegel, S., Ramos, B.M.C., 2002. Applying laboratory research: drug anticipation and the treatment of drug addiction. Experimental and Clinical Psychopharmacology 10, 162–183.

Sinha, R., Lacadie, C., Skudlarski, P., Fulbright, R.K., Rounsville, B.J., Kosten, T.R., Wexler, B.E., 2005. Neural activity associated with stress-induced cocaine craving: a functional magnetic resonance imaging study. Psychopharmacology 183, 171–180.

Venzia, P., 2007. Sensitization, drug addiction and psychopathology in animals and humans. Progress in Neuro-Psychopharmacology and Biological Psychiatry 15, 1553–1555.

Vuchinich, R.E., Heather, R. (Eds.), 2003. Choice, Behavioral Economics, and Addiction. Elsevier, Oxford, UK.

Relevant Websites

http://www.niaaa.nih.gov/Pages/default.aspx – National Institute on Alcohol Abuse and Alcoholism.

http://www.nida.nih.gov/nidahome.html – National Institute on Drug Abuse.

http://www.samhsa.gov – Substance Abuse and Mental Health Services Administration.

CHAPTER

19

Overdose

Chuan-Yu Chen[*,$], *Wei J. Chen*[**]

[*]Institute of Public Health, National Yang-Ming University, Taiwan

[$]Division of Mental Health and Addiction Medicine Research, National Health Research Institutes, Taiwan

[**]Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taiwan

OUTLINE

DEFINITION OF OVERDOSE

Definitions of overdose encompass a number of facets. One commonly cited definition is built on the intuitively straightforward idea of "excessive use," as seen in "an accidental or deliberate dose of a medication or street drug that is in excess of what is normally used" in Medical Subjective Heading (Mesh) system. However, it is important to know that although overdose has been one of major issues concerning health professionals and scientists in addiction for decades, the term of "overdose" has not been introduced into the Mesh system under the disease category of substance use disorders until 1990; between 1966 and 1989, such condition was only indexed under the term "poisoning" to describe "a condition or physical state produced by the ingestion, injection, inhalation of or exposure to a deleterious agent." The present description in Mesh, underscoring a rather obvious, yet probably distant cause, may reflect gradual evolution in conceptual issues for overdose.

With the accumulation of scientific knowledge in the past 20 years, now there seems to be a general understanding that certain individual, pharmacological, and contextual factors may also facilitate or exacerbate the risk of overdose among drug users, even when the amount of drug consumed is normal as compared with that in other drug users. That is, the excessive consumption appears to be not a necessary cause, and some scientists even argued whether the term overdose is a misnomer. Available evidence may favor another definition which in turn draws attention to the manifestation of adverse physical and mental effect shortly due to drug intake, as exemplified in the Worth Health Organization's lexicon for substance use problems. The emphasis on clinical presentation may help establish a working conceptual model which explicitly suggests that, excess use is certainly an important contributory factor, but the occurrence of overdose should depend on individual physical and contextual conditions and drug nature.

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00019-X

Overdose can occur accidentally or intentionally. To illustrate, unintentional overdose can take place when a person concomitantly uses opioid with other central nervous system (CNS) depressants (e.g. alcoholic beverages or sedatives-hypnotics), which together may lead to respiratory depression. Although relatively rare than accident cause, drug overdose can also take place as an intentional act. In the past decade, there has been an alarming increase in suicide cases due to psychotropic prescriptions, such as taking prescription more than medically recommended dose. Either accidental or intentional, some may have survived an overdose, and the ratio of nonfatal versus fatal, approximately ranging from 20 to 25, might differ by region and primary drug.

SYMPTOM AND DIAGNOSIS

A variety of psychotropic drugs have been linked with the occurrence of overdose, among which the most prevalently reported are opioids (primarily heroin), sedatives-hypnotics (e.g. barbiturate and benzodiazepines), prescribed opioid analgesics, ethanol, stimulants (e.g. cocaine, amphetamine, and Methylenedioxymethamphetamine (MDMA)), and club drugs (e.g. γ-Hydroxybutyrate (GHB)). On the basis of drug category, clinical symptoms of overdose were summarized in Table 19.1. Largely depending on type of drugs and overdose severity, the presentation of symptom constellation may lend certain indication for diagnosis. It should be noted that, even within the same medication category, overdose symptom presentation and risk of fatal cases can differ by compound, partly due to variation in onset time and duration of behavioral effects and elimination half-time; for example, Alprazolam and Flunitrazepam were more commonly mentioned in fatal overdose than other compounds of Benzodiazepines. Clinically, the diagnosis of opioid overdose commonly covers CNS and respiratory depression, myosis, together with the evidence of opioid use (e.g. needle tracks or soft tissue infection). For heroin overdose, a triad of abnormal mental status, depressed respiration, and miotic pupils may provide a sensitivity of 92% and a specificity of 76% for diagnosis. For sedative-hypnotic overdose, CNS and respiratory depression are key clinical features for diagnosis.

TABLE 19.1 List of Signs and Symptoms for Drugs Commonly Involved in Overdose

Drug	Symptoms
Heroin	Shallow breathing, pinpoint pupils, dry mouth, low blood pressure, weak pulse, bluish-colored nails and lips, bradycardia, constipation, drowsiness, delirium, disorientation, and coma.
Barbiturates	Shallow breathing, sluggishness, incoordination, slowness of speech, drowsiness, staggering, and coma.
Benzodiazepines	Mostly mild CNS symptoms (e.g. incoordination, ataxia, slurred speech, impaired attention and memory). Paradoxical reactions (e.g. delirium and hallucination) and GI symptoms were reported. Rarely fatal, unless in combinational use with CNS depressants.
Prescribed opioid analgesics (e.g. Oxycodone)	Shallow breathing, dizziness, slow heart rate, low blood pressure, confusion, and coma.
Methadone	Drowsiness, limpness, pinpoint pupils, hypotension, bradycardia, pulmonary edema or aspiration, depressed respiratory rate and depth, coma, and death. Risk of overdose increases while concomitant use with CNS depressants or medications interfering methadone metabolism (e.g. medications inhibiting Cytochrome P450 3A4(CYP3A4)).
Buprenorphine	Shallow breathing, cold or clammy skin, pinpoint pupils, fainting, slow heart rate, weak pulse, extreme drowsiness, and coma.
Ethanol	Confusion, vomiting, slowed breath, slurred speech, stupor, incoordination, and coma.
Cocaine/crack	Chest pain, palpitation, seizure, high blood pressure, increased heart rate, cardiac arrhythmias, paranoia, and delirium.
Amphetamine, Methamphetamine	Chest pain, palpitation, tremor, sweating, high blood pressure, agitation, hallucinations, seizures, tachycardia, arrhythmias, hyperthermia, delirium, and convulsions
Methylenedioxymethamphetamine (MDMA)	Nausea, cardiac arrhythmia, hyperthermia, cerebral hemorrhage, rhabdomyolysis, acute renal failure, and hepatotoxicity.
γ-Hydroxybutyrate (GHB)	Decreased consciousness, mild hypothermia, asymptomatic bradycardia, mild acute respiratory acidosis, and emesis.

Although each of the just-mentioned drugs alone can cause overdose, it is very common to find more than one drugs were involved in the event of overdose, which may increase difficulties in accurate diagnosis and treatment management.

TREATMENT

Contrary to popular conception, less than one-quarter of overdose took place within 3 h of drug administration. This critical period leaves an opportunity for early interventions such as calling an ambulance, seeking assistance, and timely transportation to emergency department. For accident and emergent management, severe clinical features (e.g. respiratory depression and coma) associated with opioids, heroin, and methadone overdose are usually treated by opioid antagonists (e.g. naloxone), with regular monitoring and incubation. Naloxone (Narcan) can effectively block the mu-opiate receptors and reverse the effects from an array of natural, semisynthetic, and synthetic opioids; for buprenorphine overdose management a higher dose of methadone is recommended given its higher affinity for mu-opiate receptors. Naloxone can be administered via intravenous and intramuscular injection, as well as intranasal insufflations. In addition to medical staff, there are also reports on the measures to expand the accessibility and reach of naloxone to high-risk opioid users and their peers.

Similarly, in order to attenuate clinical features of stimulants overdose, treatments involve medications with effects in (1) blockade of presynaptic catecholamine reuptake sites or postsynaptic receptors (e.g. bupropion, aripiprazole, risperidone, topiramate, and modafinil) and (2) reduction in drug availability in the CNS by antidrug antibodies or catabolism enhancers. In addition, careful monitoring is needed for seizures, cardiac arrhythmias, stroke, and pulmonary complication. As to managing sedative-hypnotic overdose, maintenance of the airway and ventilator support should be the first priority. Flumazenil, a competitive benzodiazepine receptor antagonist may be used to reverse the sedative effects of benzodiazepines, in particular short-acting ones.

EPIDEMIOLOGY

In general, epidemiological evidence of overdose has primarily been produced from some high-income countries in the North America (i.e. the United States and Canada), Europe (e.g. the United Kingdom, Italy, and Spain), and Australia, particularly that from longitudinal studies. In most cases, participants are recruited from clinical treatment, maintenance programs, or criminal justice system; in some situations, statistics were obtained from medical records, judiciary archives, and coronial reports. Although the impact of illegal drug use in Asia and Africa has gradually become problematic, to date, few countries have published studies assessing the problem nature of overdose and associated factors. The few studies that have been published are mostly descriptive and cross-sectional, and longitudinal evidence still awaits to be gathered.

According to the published literature, the profile of drugs involved in overdose events has slowly evolved over the past decades. From 1960 through 1980, the available evidence in many parts of the world was mainly focused on opioids or heroin, and starting in 1980s more attention has been paid to cocaine until recently. Later on, the evidence on amphetamine/methamphetamine, club drugs, and intravenous drug use-involved overdose has gradually emerged, and in the recent years prescriptions have appeared as one of the leading drugs reported in overdose, which is especially true in the United States.

PREVALENCE AND MORTALITY

Drug users are at risk of overdose, with nonfatal and fatal consequences. The risk estimates of overdose may vary quite widely, depending on drugs, sources of data, and region. In general, the lifetime prevalence of nonfatal overdose appears highest for opioids, with the estimates falling between 38 and 83%; roughly 10–30% opioids users have experienced nonfatal overdose in the past 12 months. The lifetime prevalence estimates of nonfatal overdose were 12–40% for cocaine. In terms of fatal overdose, the majority of the existing literature is about the case of heroin, which was estimated to be 2–4%. For other illegal drugs, although cases reports on MDMA, amphetamine/methamphetamine, or GHB-involved overdose have accumulated over the years, the evidence on prevalence estimates, both nonfatal and fatal, were generally lacking as compared with those for heroin and cocaine. A study in Australia indicated that roughly one in two GHB users have experience nonfatal overdose in their lifetimes. Recently, the hospitalizations for prescription opioids, sedatives, and tranquilizers-involved poisoning increased substantially in the United States and a further exploration has indicated the rise in the prevalence of nonfatal overdose may be due to the problematic use of benzodiazepine, methadone, and analgesics.

Overdose has been a major cause of premature death among drug users, especially opioids users. In 2000, it has been estimated that over 69 000 individuals died of opioid overdose globally. A recent review on cohort studies for heroin and other opioids users found the

proportion of deaths due to overdose ranged from 2.5% in Sweden and 7.3% in Portugal to 73.7% in Austria and 77.2% in the United Kingdom. The pooled estimate for overdose Crude mortality rate (CMR) was estimated to be 0.65 per 100 person-years (95% confidence intervals: 0.55, 0.75). For illegal drugs other than heroin and cocaine, available longitudinal studies on mortality were generally few. A study in Sweden on primarily cannabis, inhalants, Lysergic acid diethylamide (LSD), and stimulants users showed that the CMR of drug poisoning was roughly 0.0 5 per 1000 person-years. For methadone, the CMR of overdose was about 0.9 per 1000 person-years, according to studies conducted in Italy. As to prescription, a population-based observational study in West Virginia, US, reported the overall unintentional pharmaceutical overdose death rate was about 16 per 100 000 population in 2006, and the evidence of insurance enrollees in Washington, US showed that the mortality rate of prescribed opioid-related overdoses was approximately 6 deaths per 100 000 between 2004 and 2007. As to polydrug overdose, a series of studies provided evidence that alcohol and cocaine were the drugs most commonly found in heroin overdose death and the number of ploydrug-involved fatal seems to have a gradual rise over the past two decades.

RISK FACTORS

Over the past two decades, there have been many reviews and reports to summarize risk factors for nonfatal and fatal overdose, particularly for opioids. There is a broad spectrum of these suspected determinants, which can be generally grouped into individual-, drug-, circumstance-, and time-domains, on the basis of factor attributes.

Individual-level Characteristics

Male gender-related excess was generally found in all drug categories involved in overdose, including illegal drugs (e.g. heroin, cocaine, and amphetamine), prescription drugs (e.g. prescribed analgesics), and polydrugs. It should be noted that although drug users were predominately male, there is still male excess in likelihood of experiencing overdose among drug-experienced populations. As to age, the average age for fatal overdose typically fell between late twenties and early thirties for heroin, and the victims usually had used heroin for 5–10 years; the ages for nonfatal heroin overdose tended to be younger. Similar age-associated excess risks of fatal overdose also appear for cocaine and unintentional pharmaceutical overdose. In addition, longer health condition was another well-known risk factor affecting the risk of fatal and nonfatal overdose. Systematic diseases, such as liver cirrhosis, ventricular hypertrophy, bronchopneumonia, have frequently been indicated in fatal overdose. Furthermore, individual history of other substance use disorders and differential tolerance toward drug effects have been identified as predictors for overdose. In the case of heroin fatal overdose, many cases had the experience of lowered drug consumption prior to death, indicating the possibility that these individuals might be in a state of loss of tolerance to heroin. Finally, individual prior experience of nonfatal overdose has been linked with subsequent death due to overdose. Evidence from follow-up studies indicated that the risk was approximately 3–4-fold for those with two nonfatal overdose events, and was elevated to 7 for those who experienced three or more nonfatal overdose.

Drug-level Characteristics

Drug-level attributes may strongly influence the episode of overdose, including concomitant use of other drugs, route of administration, and purity, among many others. With respect to polydrug use, for heroin users, alcohol and benzodiazepines were two drugs commonly found in overdose cases, and it is hypothesized that the combined depressant effects may precipitate the occurrence of respiratory depression than a usual dose of heroin. Similar polydrug-excess risk in fatal overdose cases appears in cocaine (e.g. cocaine with heroin or alcohol), MDMA (e.g. MDMA with cocaine or amphetamine), and prescription drugs (e.g. prescribed analgesics with alcohol or antidepressants or methadone with benzodiazepines or prescribed analgesics). Route of administration has long been known as one of the most salient predictors for overdose, and route-related differences in overdose risk may depend on drug. To illustrate, for heroin the risk for overdose associated with smoking use was estimated to be 4–6%, significantly lower than that of intravenous injection (~60%). Only three to five per thousand nonfatal heroin overdose were intramuscular and subcutaneous routes-involved. Likewise, for cocaine, amphetamine, and methadone users, overdose tend to found in users by means of intravenous injection rather than inhalation or oral intake.

Circumstance-level Characteristics

Some studies have highlighted the role of circumstance in relation to overdose and fatality, although there still is rather meager evidence as compared with other three domains. Recent data collected from overdose witness suggested that overdose episodes occurring in public or abandoned building seem more likely to lead to deaths than those at homes. Witnesses calling an ambulance or taking to hospital was linked with lower odds of death,

yet the companies sometimes were reluctant or delayed to seek medical help. In addition, a series of studies in New York indicated ambient temperature was strongly correlated with the events of accidental cocaine fatal overdose or cocaine overdose mortality.

Time-level Characteristics

The risk of overdose may depend not only on one's chronological age and history of drug use (e.g. age of initiation and years of drug use), but also on recent period of abstinence. For example, evidence from the follow-up of recently released prisoners has shown a higher proportion of deaths were caused by overdose (20–60%). Their mortality rate was usually elevated within the first 2–3 weeks and 40-fold higher than those with same age and gender in the general population. Observed differences in the risk of overdose were also consistently in those (1) who were discharged from hospital after drug treatment or (2) who entered or re-entered the methadone maintenance programs; similarly, the mortality due to overdose often was elevated within the first month of change in treatment status.

All of the factors in each domain described above may be correlated with one another, and work together to shape one's risk of overdose. For example, for individual-level characteristics, a longer history of heroin or cocaine use may affect individual's liver or pulmonary function as a result of cumulative toxicological effects and infection, which may subsequently influence one's metabolic process or safety threshold toward drug effects. These individual factors, together with factors in other domains (such as concomitant drug use or reduced tolerance due to recent release from prison), may collectively increase ones' subsequent overdose risk or overdose mortality.

CONSEQUENCES AND COMPLICATIONS

In addition to mortality, nonfatal overdose may result in a wide range of health problems. In the general clinical literature and in population-based research, it is known that victims of nonfatal overdose often suffered from multiple systematic dysfunction manifested in pulmonary (e.g. pulmonary edema: 1–90%), cardiac (e.g. endocarditis), muscular (e.g. rhabdomyolysis), and neurological (e.g. cognitive impairment) systems. These above mentioned health problems may together influence drug users' vulnerability toward subsequent overdose mortality. For instance nonfatal overdose-related damage in neurological system may impair drug users' decision making concerning the ability to judge dose and tolerance, which may partly account for the observation that fatal overdose was more likely to be found in the older or experienced rather than the young or inexperienced.

PREVENTION

Much research effort has focused on the prevention of illegal drug overdose (particularly heroin and intravenous injected drugs) and several promising approaches to reduce illegal drug overdose and associated mortality and morbidity have been devised in clinical and community settings. For example, the enrollment in opioids substitution therapies such as Methadone and Buprenorphine programs may have effects in reducing deaths associated with overdose among heroin users. Other forms of harm-reduction activities were also found to have beneficial effects on overdose problems in drug-using populations, including need exchange program, supervised safe injection facility, and heroin maintenance. Improving responses on overdose are aimed at heroin users, families, and peers and usually involve basic training on cardiopulmonary resuscitation skills and general information of reaching medical assistance. Also, the provision of prescribed Naloxone, an opioids antagonist to reverse the effects of acute narcosis, for home use has been shown to provide significant protection from fatal overdose. Expanded access to Naloxone, including peer-administered Naloxone and providing Naloxone in prison and on release, are now being considered to offer as future preventive strategies.

Although the details about possible risk factors and mechanisms responsible for recent rise in prescription overdose mortality still need further investigation, there are some preventive efforts directed to medical professionals and patients. Examples of such preventive programs on medical professionals include continued education on physicians regarding evidence-based guidelines for opioid prescribing and provider detailing. In addition, researchers and practitioners are pursuing other pathways for preventing prescription overdose, such as the prescription drug monitoring programs and community-based health education. To date, limited research has shown that the above-mentioned approaches may produce significant changes in prescription overdose and associated mortality and morbidity.

SUMMARY

Drug overdose has long been one of the major causes for premature death in many parts of the world, especially in developed countries. Over the past two decades, the expansion of overdose has broken the social boundaries and gradually touched many facets of societies; meanwhile, clinical manifestations of overdose

sequelae have become more complicated than ever. Factors contributing to fatal and nonfatal overdose can be generally classified into four domains: individual, drug, circumstance, and time, and all of the factors in each domain described above may be correlated with one other, and work together to shape one's risk of overdose and subsequent health consequences. Clinical and community-based preventive strategies have emerged to reduce overdose mortality and morbidities, including increasing availability toward emergency health care services, providing emergency services and take-home Naloxone, and establishing prescription drug monitoring programs. Although a great number of studies have characterized the pattern and identified many factors that contribute to fatal and nonfatal overdose, literatures on overdose in drug addiction has mostly focused on opioids or heroin; much work remains to be done, especially for middle economies. For devising preventive strategies, detailed studies are needed in identifying high-risk groups, factors increasing risk of overdose, and barriers accessing overdose management and prevention, with urgent attention to polydrug and prescription; evidence-based effectiveness evaluation for preventive programs should be a priority.

SEE ALSO

Alcohol Use Disorders, Heroin Addiction, Prescription and Over-the-Counter Medications, Costs and Consequences (Morbidity and Mortality) Associated with Adolescent and College Drinking and Related Problems

List of Abbreviations

CMR	crude mortality rate
CNS	central nervous system
GHB	γ-hydroxybutyrate
MDMA	methylenedioxymethamphetamine
Mesh	Medical Subjective Heading

Glossary

Bradycardia a slow heart rate, usually defined as less than 60 beats per minute.

Rhabdomyolysis a condition in which skeletal muscle cells breakdown, releasing myoglobin together with enzymes and electrolytes from inside the muscle cells.

Relevant Websites

http://www.emedicinehealth.com/script/main/hp.asp – eMedicine Health.

http://www.harmreduction.org/section.php?id=51 – Harm reduction coalition.

http://www.health.state.ny.us/diseases/aids/harm_reduction/opioidprevention/index.htm – Opioid Overdose Prevention, New York State Department of Health.

Further Reading

Bohnert, A.S., Prescott, M.R., Vlahov, D., Tardiff, K.J., Galea, S., 2010. Ambient temperature and risk of death from accidental drug overdose in New York City, 1990–2006. Addiction 105, 1049–1054.

Burgess, C., O'Donohoe, A., Gill, M., 2000. Agony and ecstasy: a review of MDMA effects and toxicity. European Psychiatry 15, 287–294.

Chin, R.L., Sporer, K.A., Cullison, B., Dyer, J.E., Wu, T.D., 1998. Clinical course of gamma-hydroxybutyrate overdose. Annals of Emergency Medicine 31, 716–722.

Coffin, P., Sherman, S., Curtis, M., 2010. Underestimated and overlooked: a global review of drug overdose and overdose prevention. In: Cook, C. (Ed.), The Global State of Harm Reduction 2010. International Harm Reduction Association, London, pp. 113–119.

Darke, S., Hall, W., 2003. Heroin overdose: research and evidence based intervention. Journal of Urban Health 80, 189–200.

Darke, S., Zador, D., 1996. Fatal heroin overdose: a review. Addiction 91, 1765–1772.

Darke, S., Degenhardt, L., Mattick, R., 2007. Mortality amongst Illicit Drugs Users: Epidemiology, Causes, and Intervention. Cambridge University Press, Cambridge.

Degenhardt, L., Hall, W., Warner-Smith, M., Lynskey, M., 2004. Chapter 13: Illicit drug use. In: Ezzati, M., Lopez, A.D., Rodgers, A., Murray, C.J. (Eds.), Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors. World Health Organization, Geneva, pp. 1109–1176.

Degenhardt, L., Bucello, C., Mathers, B., Briegleb, C., Ali, H., et al., 2011. Mortality among regular or 8dependent users of heroin and other opioids: a systematic review and meta-analysis of cohort studies. Addiction 106, 32–51.

Galanter, M., Kleber, H.D., 2008. Textbook of Substance Abuse Treatment. The American Psychiatric Publishing, Arlington, VA.

Gaudreault, P., Guay, J., Thivierge, R.L., Verdy, I., 1991. Benzodiazepine poisoning. Clinical and pharmacological considerations and treatment. Drug Safety 6, 247–265.

Hall, A.J., Logan, J.E., Toblin, R.L., Kaplan, J.A., Kraner, J.C., et al., 2008. Patterns of abuse among unintentional pharmaceutical overdose fatalities. Journal of American Medical Association 300, 2613–2620.

Kim, D., Irwin, K.S., Khoshnood, K., 2009. Expanded access to Naloxone: options for clinical response to the epidemic of opiod overdose mortality. American Journal of Public Health 99, 402–407.

Leach, D., Oliver, P., 2011. Drug-related death following release from prison: a brief review of the literature with recommendations for practice. Current Drug Abuse Reviews Aug 12. [Epub ahead of print].

Paulozzi, L.J., Weisler, R.H., Patkar, A.A., 2011. A national epidemic of unintentional prescription opioid overdose deaths: how physicians can help control it. Journal of Clinical Psychiatry 72, 589–592.

Sporer, K.A., 1999. Acute heroin overdose. Annual of Internal Medicine 130, 584–589.

Stoové, M.A., Dietze, P.M., Jolley, D., 2009. Overdose deaths following previous non-fatal heroin overdose: record linkage of ambulance attendance and death registry data. Drug and Alcohol Review 28, 347–352.

Warner-Smith, M., Darke, S., Lynskey, M., Hall, W., 2001. Heroin overdose: causes and consequences. Addiction 96, 1113–1125.

White, J.M., Irivine, R.J., 1999. Mechanisms of fatal opioid overdose. Addiction 94, 961–972.

CHAPTER

20

Prenatal Exposure to Alcohol and Illicit Substances

Rebecca L. Schacht

Johns Hopkins University School of Medicine, Baltimore, MD, USA

OUTLINE

INTRODUCTION

Recent estimates of illicit drug use among pregnant women in the United States range from 5 to 12%. National US surveys indicate that 8% of pregnant women report illicit drug use in their first trimester of pregnancy, 4% in the second trimester, and 2% in the third trimester. A similar pattern is evident for alcohol use: 22% report first-trimester use, 7% second-trimester use, and 5% third-trimester use. Rates vary depending on substance (Table 20.1). Rates of prenatal substance use are likely to be higher in adolescent than adult women, as adolescents report higher rates of substance use in general.

Rates of prenatal substance use can be difficult to accurately assess based on patient self-disclosure, given the stigma associated with such use. Some researchers ask individual patients to report their drug use while collecting anonymous toxicology data on the patients as a group. Data from such research suggest that between 66 and 82% of prenatal use is identified via interview techniques, implying that up to one-third of prenatal substance use goes unreported. Some pregnancies may spontaneously abort before being detected or reported as a result of maternal substance use. Therefore, prenatal substance use may be underestimated. In addition, isolating the effects of specific substances can be challenging because substance use is associated with numerous other high-risk phenomena (e.g. poly-substance use, poor nutrition, violence exposure, lack of prenatal care).

Perhaps the most challenging methodological issue in regards to research on prenatal substance exposure is the problem of isolating the cause of differences found between exposed and nonexposed groups. Prenatal substance use rarely occurs in the absence of other potentially harmful environmental phenomena. Extensive research has identified numerous correlates of

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00020-6

TABLE 20.1 Rates of Substance Use in Last Month Reported by Women in 2003–04

	Pregnant	Non pregnant
Substance	**% (*n* in thousands)**	**% (*n*)**
Binge alcohol use or heavy drinking	4.5 (114)	23.3 (13 708)
Marijuana/hashish	3.6 (92)	7.5 (4397)
Crack/cocaine	0.3 (8)	1.0 (617)
Heroin	0.1 (2)	0.1 (31)
Hallucinogens	0.2 (5)	0.5 (287)
Inhalants	0.1 (2)	0.2 (126)
Nonmedical use of prescription pain relievers	1.2 (30)	2.8 (1645)
Nonmedical use of prescription stimulants	0.3 (8)	0.9 (503)

Source: Substance Abuse and Mental Health Services Administration, Office of Applied Research (2009). The NSDUH Report: substance use among women during pregnancy and following childbirth. *Rockville, MD.*

substance use in pregnancy. These include historical and current abuse, poverty, substance-using partners, a past history of substance use disorders, higher rates of sexually transmitted infections, elevated rates of serious physical illness such as HIV/AIDS and hepatitis C, and polysubstance use. Prenatal substance use is also associated with poor nutrition in pregnancy, prenatal complications such as bed rest or bleeding, preterm labor, placental insufficiency, intrauterine fetal death, low birth weight, lack of prenatal care, high blood pressure, and complication of pain management in labor. Substance-using pregnant women often meet criteria for psychiatric diagnoses in addition to their substance abuse diagnosis.

Early conceptualizations of problems associated with preterm exposure were based strictly on the physiological effects of substances. However, subsequent research has supported a model of bidirectional relationships among psychosocial processes that co-occur with substance exposure and the direct effects of the exposure. Thus, current conceptualizations consider the effects of prenatal exposure within the context in which the exposure occurs. In one study of relative risk contributions for low birth rate, the researchers assessed numerous variables, including low socioeconomic status, psychosocial stress and mental health functioning, lack of prenatal care and nutrition, and polysubstance use. Results indicated that no single substance was significantly associated with low birth weight when psychosocial and biological factors included in the statistical model. This finding suggests that prenatal drug use may function as an indicator of other problems associated with deleterious fetal outcomes, and may be one of many causes of postnatal problems, rather than the sole cause.

Another confounding factor is shared maternal–fetal genetics, which represent an additional contributor to any behavioral, structural, or neurological anomalies that may be identified in the offspring. All of these factors interact with each other, thereby further complicating the relations between prenatal substance exposure and postnatal outcomes.

In this chapter, we briefly review research on the effects of *in utero* exposure to alcohol and illicit substances, including marijuana, cocaine, methamphetamines, opioids, benzodiazepines, and inhalants. We limit our discussion to human findings, as a review of animal research is beyond the scope of this chapter.

BIOLOGICAL EFFECTS OF PRENATAL SUBSTANCE USE

Prenatal substance exposure is associated with a range of problematic outcomes in children. These include congenital malformations, low birth weight, intrauterine death, infant mortality and prematurity, fetal growth restriction, neonatal withdrawal symptoms, and lower Apgar scores. Prenatal substance exposure has also been linked with longer-term problems, such as intellectual delays, inability to focus, impulsivity, and higher rates of behavioral problems.

Effects of substance exposure can vary significantly based on the type of substance used, concurrent exposure to multiple substances, dose and frequency, and weeks of gestation. For example, an infant chronically exposed to an opioid (e.g. methadone) may be physiologically dependent at birth and experience withdrawal symptoms, but long-term prognoses can be good. In contrast, an infant exposed to moderate levels of alcohol while *in utero* might experience lifelong developmental issues as a result. Following birth, an infant's prognosis is further influenced by the nature of the care the child receives.

Fetuses are thought to be approximately 100 times more sensitive to environmental toxins than adult humans. As a result, substances that may cause little or no maternal harm may cause significant harm to the fetus. This sensitivity is due to several factors, including the fetus's small size, the speed and complexity of prenatal development, and the cascading nature of fetal development (i.e. each step of development depends on the preceding step).

Effects of substances on fetuses depend on the timing of use and the substance used. Cell development (cytogenesis) and tissue development (histogenesis) occur

mostly during the first half of pregnancy; therefore, exposure during early pregnancy is likely to affect these processes. The second half of pregnancy is dominated by cell specialization (differentiation) processes and growth. For example, a well-known problem that develops early in pregnancy is neural tube malformation, which is one of the most commonly observed congenital defects. The neural tube forms very early in pregnancy, during gestational days 23–26, often before women are aware of their pregnancy. Neural tube defects, which result in severe and often terminal developmental problems, occur when the neural tube, which evolves into the brain and spinal cord later in pregnancy, fails to close properly.

Most substances are believed to cross the placenta and exert effects on fetal well-being. However, some substances cross the placenta more readily than others, resulting in higher fetal exposure to these substances. In addition, qualities of the placenta are believed to affect the dose received by the fetus, as biological analyses of neonatal substance levels in identical twins – who share a placenta – are similar, whereas those of fraternal twins diverge. Another factor that determines the extent of a substance's effect is the sensitivity of fetal biological structures, which varies from substance to substance.

Evidence suggests that an individual's experience *in utero* affects not only that individual's health, but also the health of that individual's offspring. "Fetal programming" refers to how the uterine environment affects, or "programs," fetal development. The idea that the uterine environment could affect subsequent generations developed from observations of pregnant Dutch women exposed to severe malnutrition in 1944 as a result of World War II, and the discovery that these women's children and grandchildren are of smaller-than-average size. Subsequent research has supported the possibility that fetal programming could affect an individual throughout life and also affect that individual's offspring. Such findings suggest that addressing prenatal substance use can have cascading effects throughout subsequent generations.

"Neonatal abstinence syndrome" (NAS) refers to symptoms in newborns of withdrawal from substances as a result of prenatal exposure resulting in physical dependence. Neonates experiencing NAS symptoms can be difficult to soothe, have difficulty feeding, and exhibit neurobehavioral symptoms such as increased startle reflex. Marijuana and cocaine are believed not to cause NAS symptoms. Infants will often show signs of NAS following alcohol, opiate, and sedative exposure, but pharmacologic treatment is often either not available or not clinically warranted based on the severity of symptoms. NAS can affect the infant–caregiver relationship, as infants with NAS are challenging to care for (e.g. are difficult to soothe, do not sleep regularly, and may experience difficulties in feeding).

SPECIFIC SUBSTANCES

Alcohol

"Fetal alcohol spectrum disorders" (FASD) refers to the range of abnormalities that can result from maternal alcohol consumption. FASD does not constitute a diagnosis. "Fetal alcohol syndrome" (FAS) refers to a specific diagnosis, the criteria for which an individual on the FASD spectrum may meet. Prevalence estimates suggest that FAS occurs in 0.2–2.0 infants per 1000 born annually in the United States, although these data likely underestimate the scope of the problem because maternal alcohol use is underidentified and underreported. In addition, recent research suggests that neurobiological damage can occur without the presence of readily identifiable facial dysmorphisms typically seen in FAS.

FAS has been recognized perhaps as long as alcohol has been consumed, but did not become a named syndrome until a 1973 report in *The Lancet.* In 2004, the US Centers for Disease Control and Prevention published formal diagnostic criteria for FAS. FAS is not included in the DSM-IV-TR, but can be found in ICD-9 and ICD-10. Symptoms fall into three clusters. Facial abnormalities include a smooth philtrum (the flattened area between nose and mouth, which is ridged in typical development), thin vermillion border (upper lip), and small palpebral fissures (decreased distance between upper and lower eyelids). A second symptom cluster is inhibited growth, defined as prenatal or postnatal height and weight below the 10th percentile of typical development. The third and most diverse symptom area is central nervous system abnormalities, which includes structural, neurological, and functional damage. The effects of FAS may be most pronounced in the central nervous system, and FAS is a significant cause of intellectual disabilities. To receive a diagnosis of FAS, individual must have documented abnormalities in all of the three symptom clusters. FAS diagnostic criteria do not require confirmed prenatal alcohol exposure. The 4-Digit Diagnostic Code, developed at the University of Washington, is one commonly used diagnostic approach for suspected FAS.

Alcohol is a central nervous system depressant which crosses the placenta. It is believed to induce teratogenic effects via a variety of mechanisms, such as fetal hypoxia and oxidative stress, and interference with neuronal development. Fetal exposure to alcohol affects most brain systems, as evidenced by differences in brain structure, metabolism, and functioning. Neuroimaging studies indicate that these differences persist at least

through adolescence and, in some cases, are present even in children without facial dysmorphism.

Some research suggests that even modest quantities of alcohol – seven or fewer drinks per week, or three or more drinks in one drinking episode – can have teratogenic effects. Other research suggests a threshold of 15 drinks per weeks before fetal harm occurs. The reasons for these discrepant findings are not entirely clear. It is possible that the effects of alcohol exposure on a developing fetus vary widely among individuals for reasons that have not been identified. Without an established safe dose and little evidence to suggest that alcohol consumption confers significant health benefits in pregnancy, women have long been advised to maintain abstinence from alcohol throughout their pregnancies. However, approximately 13% of pregnant women consume some alcohol during their pregnancy. Rates of prenatal alcohol consumption are higher in at-risk populations, including individuals who have been treated for alcohol use or mental health disorders, tobacco smokers, economically marginalized populations, and women experiencing physical abuse or other forms of stress. Data from the Pregnancy Risk Assessment Monitoring System (PRAMS) indicate that higher income, older age (35+ years), and years of education are also linked with prenatal alcohol use.

Alcohol consumption can be a particularly problematic substance in pregnancy for several reasons. First, maternal alcohol consumption confers risks of severe teratogenic effects for the fetus. Second, the fetus is believed to be particularly vulnerable to alcohol's effects very early in pregnancy – often before a woman knows that she is pregnant. Third, alcohol is widely available, widely consumed, and legal, which may imply to the general population that it is safer than other substances. Fourth, social norms vary widely in regards to acceptable levels of consumption, particularly for younger drinkers (*see* Peer Influences on Addiction). Fifth, drinkers' estimates of their consumption are often based on number of glasses or number of drinks consumed, which are often not reflective of the actual number of ounces consumed. Drinkers' memories may also be impaired by alcohol's effects on cognitive function. Therefore, women may underestimate the amount of alcohol they consume and thus underestimate the risks to their fetuses. For these reasons, pregnant women are advised to maintain complete abstinence from alcohol, particularly in the first trimester of pregnancy.

FAS can have lifelong consequences for offspring, and fetal development is less severely affected the earlier in the pregnancy the alcohol exposure stops. Therefore, early identification of prenatal alcohol consumption is vital. Nevertheless, stopping even late in pregnancy protects fetuses from subsequent harm. Mothers should be encouraged throughout pregnancy to become abstinent from alcohol.

Marijuana

Marijuana is the most commonly used illicit substance by pregnant women. Approximately one-third of the psychoactive component of marijuana – Tetrahydrocannabinol (THC) – crosses the placental barrier, thereby exposing the fetus to significant doses of THC. In spite of marijuana's popularity, fetal exposure to the active ingredient in marijuana remains understudied and questions remain regarding its effects on developing fetuses.

Much of the data regarding longitudinal outcomes associated with prenatal marijuana exposure are from two epidemiological studies: the Ottawa Prenatal Prospective Study (OPPS) and the Maternal Health Practices and Child Development Study (MHPCDS). In these studies, marijuana exposure has not been associated with increased risk of miscarriage or other perinatal complications, but has been associated with low birth weight. Marijuana is typically consumed via smoking, which exposes the fetus to carbon monoxide. Low birth weight, which has been associated with prenatal cannabis use in these studies as well as in meta-analytic studies, is thought to be the result of carbon monoxide exposure.

Research suggests that prenatal cannabis exposure is associated with developmental delays observable at birth. Findings from the OPPS and MHPCDS are consistent in implicating marijuana exposure in higher rates of increased tremors and exaggerated startle reflex in neonates. Alterations in neonates' visual functioning have also been reported. These delays appear to be mild and to disappear by the age of 6 years. Based on the research to date, THC appears to be less likely than other substances to cause congenital malformations in fetuses, although findings are not entirely consistent.

Much of the research assessing differences following prenatal cannabis exposure focuses on cognitive ability. Evaluations of cognitive ability in older cannabis-exposed children have revealed mixed results. Some studies suggest that prenatal marijuana exposure is associated with lower intelligence quotient, whereas others reveal no effect. Similar contradictory patterns of findings have been discovered for other forms of frontal lobe functioning, such as planning and perception, with some studies showing significant associations and some not. Prenatal cannabis exposure also appears to be associated with higher levels of antisocial or delinquent behavior in late childhood and adolescence. In addition, cannabis exposure has been implicated in disruption of neuronal development in the fetal brain, particularly in areas relating to mood, executive function, and reward. One explanation for these apparently

contradictory findings is that prenatal cannabis exposure affects some types of executive function while leaving others intact. Although the extant research suggests that prenatal cannabis exposure is associated with developmental issues, behavioral differences following prenatal cannabis exposure could also be accounted for by other environmental factors (e.g. polydrug exposure, parenting approach, and socioeconomic status). Additional research is needed to more precisely identify areas of concern following prenatal cannabis exposure.

Cocaine

Approximately 15–17% of cocaine users are women of childbearing age. Cocaine is a central nervous system stimulant that increases catecholamine ("fight or flight" hormone) levels, thereby activating the sympathetic nervous system. As a vasoconstrictor, cocaine increases maternal blood pressure, thereby increasing risk of hemorrhage and decreasing uterine blood flow and fetal oxygen supply. Cocaine also appears to block reuptake of some neurotransmitters, including serotonin, dopamine, and norepinephrine. Cocaine is believed to readily cross the placenta, thereby directly exposing the fetus to its effects. However, analysis of neonates' hair suggests that fetuses do not receive the entire dose of cocaine ingested by the mother.

Initial reports on cocaine-exposed children ("crack babies") in the 1980s and 1990s indicated that prenatal cocaine exposure was associated with severe teratogenic effects. More recent research suggests that prenatal cocaine use often occurs in tandem with other problems, and many of the effects initially attributed to prenatal cocaine exposure are at least partially attributable to other comorbid phenomena (e.g. poor nutritional, maternal stress and psychiatric illness, polysubstance use, lack of prenatal care, chronic medical problems such as HIV/AIDS).

However, researchers have discovered that some problematic effects associated with prenatal cocaine exposure persist independently of those resulting from other teratogens. These effects, which intensify as doses increase, include higher rates of preterm labor and early birth, low birth weight, shorter gestational period, placental abruption (placental separation from uterine wall), and neonates who are small for their gestational age.

Evidence suggests that prenatal cocaine exposure is also associated with functional deficits in later childhood. Children who are prenatally exposed to cocaine tend to exhibit small but detectable deficits in academic, intellectual, and language functioning relative to nonexposed children. In addition, neurocognitive deficits in maintaining attention and regulating behavior have been observed in cocaine-exposed children. These deficits appear to result from inhibited neurological developmental processes; as a result, they may not be evident until children reach school age and attempt tasks that require increasingly complex brain function. Most researchers who evaluate postnatal effects of cocaine exposure use demographic and other environmental factors as statistical covariates, thus increasing their capacity to isolate effects of cocaine exposure from those of other environmental factors. However, covariates and other research methods differ from study to study. As a result, the task of identifying which observed effects are due to prenatal exposure and which effects are due to other factors is ongoing.

Opioids

Babies exposed to opioids, including heroin, methadone, and prescription drugs (e.g. OxyContin, Percocet, and Vicodin) while *in utero* often experience withdrawal symptoms after birth. Although rarely fatal, opioid withdrawal can cause significant infant distress. Withdrawal is evident in four systems: autonomic nervous (e.g. sleep disturbance, fever, tearing and sweating, disrupted sucking patterns), central nervous (e.g. tremors, seizures, twitching, irritability, hyper- or hypo-activity, high-pitched crying), respiratory (nasal congestion, sneezing, respiratory distress), and gastrointestinal (e.g. vomiting, diarrhea, losing or failing to gain weight, hiccups). Withdrawal symptoms appear in 55–94% of neonates with prenatal opioid exposure. Signs of withdrawal appear within several minutes to 72 h after birth, depending on the mother's last opioid ingestion and amount, fetal health and gestational age, and type of analgesia used during labor. No clear dose–response relationship between amount of exposure and severity of withdrawal symptoms has been established.

Infants typically recover from prenatal opioid exposure after initial withdrawal symptoms resolve, but some long-term developmental problems in opioid-exposed children have been observed, including higher rates of attention deficit hyperactivity disorder, impulsivity, and diminished motor and language performance compared to nonexposed children. However, the research in this area is scant and additional work is necessary to draw firm conclusions regarding prenatal opioid exposure and long-term prognoses in children.

Methamphetamines

After marijuana, methamphetamines are the most commonly used illicit substance in pregnancy. Comparisons of methamphetamine levels in neonatal and maternal hair indicate that the levels are comparable, implying that methamphetamine readily crosses the

placenta without any decrease in dose. Methamphetamine, like cocaine, is a central nervous system stimulant with vasoconstrictive properties, and thus confers similar risks as cocaine such as fetal hypoxia and decreased uterine blood flow. Like cocaine, methamphetamine alters neurotransmitter levels, such as increasing levels of serotonin, dopamine, norepinephrine, and amines, while increasing catecholamine levels. Methamphetamine may also decrease maternal appetite, leading to poor fetal nutrition and, as a result, intrauterine growth restriction. Prenatal methamphetamine use has been associated with other pregnancy complications such as preterm birth and placental abruption.

Methamphetamine-exposed neonates tend to be smaller than their nonexposed peers. In addition, they may exhibit signs of exposure that mimic those observed in cocaine-exposed neonates, such as increased startle response due to hyperarousal. Birth defects such as cleft palate and cardiac anomalies have been observed in methamphetamine-exposed neonates. Methamphetamine-exposed neonates have been found to exhibit more severe health problems, such as fetal distress and neurological and physiological anomalies, relative to opioid-exposed neonates. Smaller head circumference, higher rates of behavioral problems, and lower academic functioning have been observed in older children exposed to methamphetamine prenatally. Neurocognitive anomalies in childhood have also been observed, including decreased attention span and working memory capacity. These phenomena have been observed in some samples, but not all, possibly as a result of methodological variability between studies. In addition, as with other assessments of prenatal outcomes following substance exposure, the behavioral, structural, and neurological anomalies that are observed could also be the result of problems that occur concurrently with methamphetamine exposure.

MDMA (3,4-methylenedioxymethamphetamine; Ecstasy) is structurally similar to methamphetamine. The effects of prenatal MDMA exposure in humans have not been widely studied. Limited extant data suggest that prenatal MDMA exposure is associated with higher rates of cardiac malformations and clubfoot in infants.

Inhalants

Inhalant abuse includes inhalation of household products containing toluene, which is an organic solvent that readily crosses the placenta. Products containing toluene include paint, gasoline, ink, lighter fluid, and glue. Other abused inhalants include nitrous oxide (laughing gas), chloroform, and ether; however, toluene is currently the most widely abused inhalant in the United States. Symptoms of prenatal toluene exposure in neonates are referred to collectively as fetal solvent syndrome or toluene embryopathy. Symptoms include early parturition, fetal growth restriction, and facial dysmorphisms similar to those observed in FAS. NAS following toluene exposure has been observed and is characterized by excessive crying, tremor and increased Moro reflex, insomnia, and poor feeding. Developmental delays in later childhood have also been observed. Although additional research is necessary to isolate the effects of toluene exposure from other forms of prenatal substance exposure, toluene effects appear to occur reliably at high doses, suggesting that at least some fetal changes following toluene exposure can be attributed to the substance and not to other substances or environmental exposures.

Benzodiazepines

Benzodiazepines, including diazepam (Valium), alprazolam (Xanax), and clonazepam (Klonopin), are commonly misused prescription medications. Several reviews and meta-analyses have been conducted to evaluate potential teratogenic effects of benzodiazepine medications, including diazepam. Some data suggest that benzodiazepine exposure is associated with higher rates of cleft palate or other oral malformations in the fetus. However, this finding is not consistent across samples, and despite repeated attempts to establish a definitive answer, there is currently no consensus regarding the safety of benzodiazepines in pregnancy. Until such consensus is achieved, pregnant women should be advised to avoid licit and illicit use of benzodiazepines.

Phencyclidine

Research on the effects of prenatal phencyclidine (PCP; also known as angel dust) exposure in humans is limited. However, reports suggest that *in utero* PCP exposure is associated with signs of neonatal PCP intoxication (e.g. jitteriness, irritability, and digestive issues). In addition, PCP may result in birth complications including early parturition and low birth weight. Limited research suggests that, when other drug use is statistically controlled for, prenatal PCP exposure contributes independently to compromised neonatal functioning. However, it remains unknown whether PCP exerts effects on children after they are beyond the neonatal stage.

ADDRESSING PRENATAL SUBSTANCE USE

Detection

All women of childbearing age should be screened for substance use, particularly those who are sexually active

but do not consistently use a reliable form of birth control. Some research indicates that health care providers feel unprepared to address possible patient resistance should the provider broach the topic of substance abuse treatment. Lack of referral options can also be a limitation. However, as approximately one-half of pregnancies in the United States are unplanned, universal screening is vital.

The simplest means of detecting prenatal substance use is by asking the patient about her substance consumption. This approach is complicated by stigma and maternal concerns of possible legal involvement, and the health care provider should be aware that these concerns may result in reporting bias. Thus, health care providers are advised to ask open-ended questions and maintain a nonjudgmental demeanor with patients. This approach is likely to maximize the potential for honest responding and is consistent with The American College of Obstetricians and Gynecologists' (ACOG) statement, which suggests that addiction be treated not as a "moral failing" but as a chronically relapsing condition that should be treated medically. The T-ACE and the TWEAK are brief (4–5 items) alcohol screening questionnaires, similar to the CAGE, which are validated for use with pregnant women, but are limited in that their focus is on assessing alcohol consumption only. Additional assessment via interview or questionnaire should be conducted to identify use of substances other than alcohol. As prenatal substance use rarely occurs without additional concurrent psychosocial problems, additional assessment of the psychosocial context is advisable.

Biological assessment – typically using blood or urine samples – is complicated by the short half-lives of some substances (e.g. alcohol), which allows only a brief window during which a substance is biologically detectable in a woman's body. Cocaine and alcohol are rendered undetectable in fetal blood and urine after only one week, meaning that testing neonates would identify only use that occurred during the week before birth. Longer-lasting biological markers of prenatal substance use and exposure include hair cells (maternal and neonatal) and fetal meconium (first stool). Infants' hair grows during the last trimester of pregnancy; therefore, any exposure identified based on neonatal hair samples must have occurred during this period.

Education and Prevention

Public health campaigns in recent decades regarding the risks of prenatal substance use appear to have been successful, as substance use in pregnancy has dropped. However, serious and potentially widespread risks remain. Given the high rates of unplanned pregnancies, many women consume substances before discovering that they are pregnant. In addition, many women may underestimate the risks of prenatal substance use for their fetuses. Substance Abuse and Mental Health Services Administration (SAMHSA) survey findings indicate that over 80% of women who meet criteria for substance use disorders do not desire treatment or believe that treatment is necessary, implying that additional public education about substance use disorders may be necessary. Therefore, educating young women about the risks of prenatal substance use, particularly early in pregnancy, is vital. This information can be disseminated broadly, such as via public health campaigns, or person-to-person, such as by health care providers. Indeed, ACOG suggests that health care professionals provide women with psychoeducation and referrals to substance abuse treatments.

Incarceration and Other Forensic Involvement

Maternal substance use in pregnancy is defined as child abuse in 15 US states and is grounds for civil commitment in 3 states. In 14 states, health care providers are legally bound to report maternal prenatal substance use and, in 4 states, to test for prenatal substance exposure when it is suspected.

ACOG issued a statement in 2011 criticizing forensic involvement in prenatal substance use. ACOG based their stance on evidence of deleterious effects that such policies can have. Legally mandated reporting of suspected prenatal substance use can result in patient mistrust of medical providers and may deter substance-using women from seeking prenatal care. Indeed, extant data suggest that prenatal substance abuse is associated with significantly lower rates of prenatal care. Lack of prenatal care is a known risk factor for birth complications, such as preterm labor, low birth weight, and mortality. Prenatal care is also an opportunity to engage pregnant women in substance abuse treatment. Thus, statutes that discourage substance-abusing women from seeking prenatal care may result in unintended consequences that further compromise the health of the fetus without resulting in decreased substance use.

As noted in ACOG's report, such policies – in addition to carrying significant risk of harm – may also be of limited benefit. Irrespective of providers' professional opinion on statutes, those involved in caring for pregnant women should know the statutes in their jurisdiction regarding prenatal substance use. In addition, providers should be ready to answer patients' questions about what is likely to occur if a report is made (e.g. incarceration, referral to treatment, Child Protective Services involvement).

Treatment

Once substance use has been identified, women should be offered treatment. Also, as substance use is associated with high rates of complications for mother and child, additional assessments might also be indicated at diagnosis and throughout pregnancy. These include HIV testing, liver and heart function assessments, and monitoring for infection. Substance use disorders are typically comorbid with Axis I and II disorders; therefore, additional psychiatric evaluation may also be indicated. Finally, pregnant substance-using women are at elevated risk for violence exposure and should be assessed and monitored accordingly. The Antenatal Psychosocial Health Assessment (ALPHA), developed at the University of Toronto, is designed to assess pregnant women's risk exposure in multiple domains: family factors (e.g. stress and support), maternal factors (e.g. extent of prenatal care, attitude toward pregnancy, current or past psychiatric problems), substance use, and current or past family violence exposure.

Brief, motivational interventions (*see* Motivational Enhancement Approaches) to engage women in treatment have been empirically supported in primary care settings, and the American College of Obstetricians and Gynecologists recommends such approaches for pregnant women. Motivational techniques are especially appropriate for pregnant women, as their concern for their babies' well-being provides an additional source of motivation to change.

Unfortunately, effective substance abuse treatment programs may not be available for pregnant women who want treatment. In addition, some women may have logistical barriers to attending such programs, such as limited transportation, childcare, and finances. Such barriers, in addition to concerns of legal involvement for women seeking treatment, may prove insurmountable for many. A 1998 study indicated that 7% of substance abuse treatment centers in the United States offer prenatal care services and 9% offer childcare services. However, a 2003 survey indicated that 59% of methadone clinics offered specialized care for pregnant women and that most (83%) consider pregnant women to be priority admissions.

Ideally, treatment for substance-using pregnant women, in addition to treating substance abuse, should comprehensively address the medical needs of mother and fetus including the treatment of comorbid psychiatric disorders. Case management is also indicated to address potential contributors to substance use such as socioeconomic, lack of vocational training and housing, legal, transportation, and childcare issues. Maternal exposure to violence should be addressed if present. Risk assessment and safety planning should be conducted with any women for whom current violence exposure is suspected. In addition to the typical stress associated with caring for a newborn, substance-using mothers may face additional challenges in parenting a fragile, substance-exposed infant with complex needs.

The rationale for a comprehensive treatment approach is strengthened by the large body of research indicating that prenatal substance exposure is almost always associated with other environmental issues. Without treatment, factors that perpetuate substance abuse during pregnancy are likely to persist following pregnancy. The postpartum period represents additional opportunities for health care providers to intervene and improve the infant's prognosis by teaching the mother parenting approaches that can attenuate the effects of substance exposure by improving the parenting relationship. For example, instruction on effective techniques for soothing a chronically fussy infant can reduce maternal stress and improve maternal functioning, including potentially reducing substance use. Research indicates that such "wrap-around care" in high-risk populations is associated with better health for neonates and significant cost reductions compared to noncomprehensive care.

Opioid Treatment in Pregnancy

Methadone treatment for opioid-dependent pregnant women has been the standard of care in the United States since 1970s. Before then, the standard of care – per FDA requirements – was that pregnant women be weaned from opioids. This standard was developed after the US FDA's requirement that women be weaned from opioids was associated with fetal death due to opioid withdrawal. Currently, federally prescribed indications for methadone maintenance include opioid dependence, per DSM-IV-TR criteria, lasting over a year (*see* Heroin Addiction).

Methadone maintenance confers benefits over abstinence- or detoxification-based approaches for multiple reasons. It increases patients' contact with health care providers because patients typically must present at a clinic daily to receive their medication. In addition, research indicates that the likelihood of relapse to illicit drug use – and its associated risks – is significantly lower with methadone maintenance than with abstinence-based treatment. Also, attrition is lower in methadone programs than for abstinence programs. Despite evidence that methadone programs are beneficial for mothers and infants, the use of methadone in pregnancy can be controversial, in part because administering methadone to the mother results in fetal exposure and, often, NAS.

Recently, researchers have begun evaluating buprenorphine as an alternative to methadone treatment for heroin addiction in pregnancy. Like methadone, prenatal buprenorphine exposure results in NAS. However, the NAS effects of buprenorphine are less pronounced than those of methadone. Switching from methadone to

buprenorphine maintenance therapy may not be indicated in pregnancy, as the mother and fetus may experience withdrawal symptoms in the process. Therefore, buprenorphine maintenance in pregnant women may be most appropriate for women who are initiating heroin replacement therapy or who are already being maintained on buprenorphine prior to conception.

Postnatally, treatment for opioid withdrawal consists of administration of diluted opium tincture (laudanum), which consists of 10% opium in an alcohol-based solution. Methadone can be problematic for use with neonates because of challenges in establishing safe and effective dosing. Swaddling the neonate in a quiet, dimly lit environment may also support the infant's recovery from withdrawal symptoms. Research to date suggests that the short- and long-term prognoses for the opioid-exposed neonate appear to be good.

Because methadone treatment for pregnant women typically is one element of comprehensive treatment services, the improved outcomes seen with methadone treatment may be due to increased access to prenatal care. Nevertheless, the benefit associated with prenatal methadone treatment for opiate dependence is clear. Pregnant women in methadone programs experience fewer pregnancy complications and health issues compared to those who continue using heroin throughout their pregnancy.

Women typically return to substance use soon after their children are born, with the majority who reported using substances pre-pregnancy or prenatally resuming use by their child's first birthday. Future work could build on the current success of public health campaigns to reduce prenatal substance consumption by expanding the message to include early childhood. In addition to reducing prenatal health risks, decreasing women's substance use perinatally also presents an opportunity to support women in continuing to make healthy choices regarding substances as they develop as parents.

For information regarding behavioral treatment of perinatal substance use, see 00174.

CONCLUSION

Prenatal substance exposure is associated with significant harm to fetuses. Substance use in pregnancy is often a symptom of other environmental stressors that may be equally harmful to offspring. Effectively treating prenatal substance use requires comprehensive care that addresses the factors that drive women's substance use. Untangling these issues in a way that can affect meaningful changes in the lives of substance-using pregnant women and their children is a complex task which, if successful, can have significant payoff for both generations, and potentially future generations as well.

List of Abbreviations

ACOG	The American College of Obstetricians and Gynecologists
FAS	Fetal alcohol syndrome
FASD	fetal alcohol spectrum disorders
MDMA	3,4-methylenedioxymethamphetamine
MHPCDS	Maternal Health Practices and Child Development Study
NAS	neonatal abstinence syndrome
OPPS	the Ottawa Prenatal Prospective Study

Glossary

Catecholamines hormones, including epinephrine (adrenaline), norepinephrine, and dopamine released by the adrenal glands in response to stress. Foster "fight or flight" reaction by increasing heart rate and respiration, dilating pupils, and increasing blood glucose.

Cytogenesis cell formation and development.

Fetal alcohol syndrome ICD-9 and ICD-10 diagnosis that describes changes in a fetus following alcohol exposure while *in utero*. Includes facial abnormalities, inhibited growth, and neurological changes associated with developmental delays.

Fetal hypoxia lack of sufficient oxygen supply to fetus that can damage central nervous system cells.

Fetal programming hypothesized association between fetal environment and health of offspring and subsequent generations. Also known as the "Barker hypothesis."

Histogenesis process by which cells differentiate into specific tissue types.

Intrauterine growth restriction slowed or restricted fetal growth while *in utero* such that fetal weight is in the 10th percentile for age. Assessed by measuring uterine fundal length (distance from mother's pubic bone to top of uterus).

Low birth weight weight of less than 2500 g (5 lb. 8 oz.) at birth.

Moro reflex an infant's involuntary response to feeling of falling characterized by reflexive spreading of arms, followed by drawing in of arms to body. Typically disappears by age 3–5 months. Absence of the Moro reflex indicates central nervous system dysfunction.

Neonate newborn baby up to 28 days old.

Neural tube defects occur when the neural tube (embryonic cells that become the brain and spinal cord) does not close properly (occurs in gestational days 23–26), resulting in a hole in the neural tube. Approximately 1 in 1000 babies in the United States are born with neural tube defects. Examples include spina bifida and anencepaly.

Parturition childbirth.

Placental abruption separation of placenta from uterine wall after 20 weeks' gestation and before birth. In severe cases, associated with maternal hemorrhage and fetal injury or death.

Placental insufficiency inability of placenta to deliver sufficient levels of nutrition and oxygen to fetus due to diminished blood flow.

Preterm birth and labor Uterine contractions that are strong and coordinated enough to cause progression of labor and possibly birth between weeks 20 and 37 of gestation. Also known as early parturition.

Teratogen substance capable of causing harm to fetus by altering development.

SEE ALSO

Families and Addiction, The Intergenerational Transference of Addiction

Further Reading

Ackerman, J.P., Riggins, T., Black, M.M., 2010. A review of the effects of prenatal cocaine exposure among school-aged children. Pediatrics 125, 554–565.

Huizink, A.C., 2009. Moderate use of alcohol, tobacco and cannabis during pregnancy: new approaches and update on research findings. Reproductive Toxicology 28, 143–151.

Jones, H.E., Tuten, M., Keyser-Marcus, L., Svikis, D.S., 2006. Specialty treatment for women. In: Strain, E.C., Stitzer, M.L. (Eds.), The Treatment of Opioid Dependence. The Johns Hopkins University Press, Baltimore, MD, pp. 455–484.

Jutras-Aswad, D., DiNieri, J.A., Harkany, T., Hurd, Y.L., 2009. Neurobiological consequences of maternal cannabis on human fetal development and its neuropsychiatric outcome. European Archives of Psychiatry and Clinical Neuroscience 259, 395–412.

National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention. 2004. Fetal alcohol syndrome: guidelines for referral and diagnosis. Available at cdc.gov/.

Ondersma, S.J., Svikis, D.S., Schuster, C.R., 2007. Computer-based brief intervention: a randomized trial with postpartum women. American Journal of Preventive Medicine 32, 231–238.

Roussotte, F., Soderberg, L., Sowell, E., 2010. Structural, metabolic, and functional brain abnormalities as a result of prenatal exposure to drugs of abuse: evidence from neuroimaging. Neuropsychological Review 20, 376–397.

Substance Abuse and Mental Health Services Administration, Office of Applied Research. 2009. The NSDUH Report: substance use among women during pregnancy and following childbirth. Rockville, MD. Available at samhsa.gov/.

Relevant Websites

http://www.cdph.ca.gov/programs/perinatalsubstanceuse/pages/default.aspx – California Department of Public Health, Perinatal Substance Use Prevention.

http://www.cdc.gov/ncbddd/index.html – National Center on Birth Defects and Developmental Disabilities (NCBDD).

http://drugabuse.gov/tib/prenatal.html – National Institute on Drug Abuse (NIDA), Prenatal Exposure to Drugs of Abuse.

http://www.ncsacw.samhsa.gov/resources/substance-exposed-infants.aspx – Substance Abuse and Mental Health Services Administration (SAMHSA), Substance-Exposed Infants.

CHAPTER

21

Impulsivity, Disinhibition, and Risk Taking in Addiction

Elizabeth K. Reynolds[*,$], *Anahi Collado-Rodriguez*[$], *Laura MacPherson*[$], *Carl Lejuez*[$]

[*]Johns Hopkins School of Medicine, Baltimore, MD, USA,
[$]University of Maryland College Park, College Park, MD, USA

OUTLINE

INTRODUCTION

Addiction involves continued engagement in a particular behavior despite adverse consequences. There are a wide range of addictive behaviors including substance use, binge eating, and gambling. Although each has unique features, there are common themes including diminished control over the problematic behavior, an appetitive urge or craving state before engagement in the problematic behavior, and a hedonic quality during the performance of the problematic behavior. For these reasons, addictive behaviors have been conceptualized as disorders of misdirected motivation and impaired self-control. A substantial body of research utilizing a variety of techniques (e.g. behavioral, neurobiological, imaging, self-report) has demonstrated a strong association between addictive behaviors and the constructs of disinhibition, impulsivity, and risk taking propensity. In the current chapter, we review these constructs including key definitions and consideration of their interrelationships, followed by a review of their relationship to addictive behaviors.

DEFINING DISINHIBITION, IMPULSIVITY, AND RISK TAKING

Disinhibition

It is important to begin the definition of disinhibition by acknowledging that it appears in the literature in two distinct ways: (1) as an umbrella term referring to a broad category of inhibitory dysfunction or behavioral undercontrol and (2) as a specific subtype/dimension of

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00021-8

impulsivity that refers to a process in which an individual has reduced capacity to alter an immediate response to some situation.

In considering the first use, disinhibition as an umbrella construct, the term is used to refer to an overarching category to include broadly overlapping but nonredundant variables such as impulsivity, sensation seeking, and risk taking propensity. As such, a range of behaviors fall into this category including loss of restraint, lack of regard for conventions, and poor risk assessment (i.e. tendency to focus on and pursue reward, even in the face of punishment). A number of reasons why these behaviors occur have been proposed including a neurobehavioral vulnerability (frontal lobes implicated), difficulty thinking about the consequences of behavior, misinterpreting social cues, poor social adjustment, inability to communicate in an appropriate way, discomfort, and response to provocation.

Disinhibition also is conceptualized as a specific type of impulsivity and is defined as manifestations of core impairments in inhibitory control, a self-regulatory process that has been conceptualized as the ability to withhold a prepotent response, interrupt an ongoing response, and protect cognitive activity from interference. This conceptualization of disinhibition as a dimension of impulsivity is measured by tasks in which participants are expected to inhibit prepotent motor behaviors or, said more simply, stop a behavior once it has been initiated. For example, the Stop and Go Task begins with the presentation of either an *X* or an *O* in the center of the computer screen. Subjects are instructed to press the "*z*" key when the *X* appears and the "/" key when the *O* appears. The letters are presented at 2-s intervals, and reaction times (RTs) are recorded. On 25% of the trials (25% of the *X* trials and 25% of the *O* trials), a tone (stop-signal) sounds after the presentation of the *X* or *O*. Subjects are instructed to refrain from pressing any keys when they hear the sound. The delay from the onset of the letter presentation to the onset of the tone (stop-signal delay) is systematically adjusted in 50-ms increments. If the subject fails to refrain from pressing a key after hearing the tone, the stop-signal delay is decreased by 50 ms on the following stop-signal trial. If the individual successfully refrains, the stop-signal delay is increased by 50 ms on the next trial. Eventually, the stop-signal delay will reach a duration at which the subject will inhibit his or her key press responses on approximately 50% of the trials. Individuals characterized by greater disinhibition are unable to inhibit their responses at longer delays (i.e. can't stop a response when the signal to stop comes later into the response) and therefore require a shorter delay to reach the point at which they are able to inhibit their responses 50% of the time. For our purposes, we will focus on this latter definition, and therefore, discuss disinhibition as a dimension of impulsivity rather than an umbrella term referring to a broad category of inhibitory dysfunction or behavioral undercontrol.

Impulsivity

While the term "impulsivity" is commonly used, it is a construct that has poor universal meaning. Several definitions of impulsivity have emerged including poorly conceived and/or prematurely expressed actions that are either inappropriate or unduly risky for a particular situation, a tendency to act without forethought (i.e. respond quickly and without reflection), excessive discounting of rewards and punishments as a function of their delay, an inability to inhibit inappropriate behavior (see definition of disinhibition above), being impatient when asked to wait, having a short attention span, and difficulty persisting at a particular activity. This multitude of definitions has led to confusion and disagreement as to the exact nature of the construct and the appropriate means of measurement. One approach to resolving the discrepancy in the definitions has been to treat impulsivity as a multidimensional construct. Despite general agreement on this matter, there remains little consensus regarding how to best characterize the different components.

A number of proposals have been set forth to define the multiple dimensions of impulsivity. First, some researchers have proposed that there are two major components of impulsivity: (1) resisting urges versus giving in to urges and (2) responding immediately to a stimulus versus planning before making a move. Second, other researchers have defined impulsivity with three subfactors: acting on the spur of the moment (motor activation), not focusing on the task at hand (attention), and not planning or thinking carefully (lack of planning). More recently, based on a comprehensive review of the literature, researchers have suggested that impulsivity is best conceptualized as a two-dimensional trait. The first trait, labeled Reward Drive, is argued to reflect individual differences in sensitivity to incentive motivation and engagement in appetitive behavior on detection of reward cues. Individual differences in reward sensitivity are thought to be mediated by the strength of responding in the mesolimbic dopamine system. The second trait, rash impulsiveness, is proposed to reflect individual differences in the ability to modify or inhibit prepotent (reward drive-initiated) behaviors in light of potential negative consequences. Individuals characterized by rash impulsiveness tend to make unwise or irrational choices because they fail to anticipate consequences or plan ahead. They may be described as insensitive to consequences, particularly when the reward sensitivity is high, leading them to perseverate in well-established (i.e. dominant)

behaviors, even when the behavior is no longer rewarded or elicits punishment. Individual differences in rash impulsiveness are thought to be mediated by the orbitofrontal cortex and anterior cingulate cortex, with the involvement of the dopamine and serotonin systems. Across the different definitions and usages, the core features seem to be (1) the tendency to execute actions too hastily or in a thoughtless manner, (2) difficulties withholding or inhibiting actions, and (3) the tendency to seek out immediate gratification at the cost of longer-term gains. Although these categories fit well with theory and empirical evidence and correspond with several widely accepted operational definitions, the exact nature of the dimensions remains controversial, and there is evidence for both relatedness and independence. More globally, the careful reader can see that although dimensions help emphasize the different aspects of impulsivity, the resulting range of dimensions and particular definitions as they appear form one conceptualization of the dimensions to another complicate efforts to develop a degree of consensus when considering what exactly is meant by impulsivity across different sources. We will attempt as we proceed here to distill the dimensions down to a manageable and useable framework, focusing on common themes across conceptualizations as opposed to sticking closely to any particular one.

Various measures have been developed to assess impulsivity, including self-report measures that rely on self-perceptions of behavior at a trait level and behavioral tasks that measure overt behavior which is thought to tap trait level measurement while also providing the opportunity to measure state responding. Three commonly used self-report measures to assess trait level impulsivity are the Barratt impulsiveness scale (BIS), the Eysenck impulsiveness scale (EIS), and the UPPS impulsive behavior scale (UPPS). The BIS is one of the most widely used self-report measures of impulsive personality traits; it has been through several revisions, with version 11 being the most recent (BIS-11). The scale consists of 30 self-descriptive items, with responses in a four-point Likert-type scale ranging from "*Rarely/Never*" to "*Almost Always/Always*." The questionnaire contains three subscales: attentional impulsiveness (actions precipitated by lack of attention), motor impulsiveness (hyperactivity due to need of movement), and nonplanning impulsiveness (attitudes and conclusions precipitated by lack of reflection). The EIS is also a widely used, self-report measure. In conceptualizing the scale, the Eysenck's distinguished impulsiveness from venturesomeness and separate scales were constructed to assess these constructs. The EIS is a 23-item, yes–no scale that is face valid for impulsivity (e.g. "Do you generally do and say things without stopping to think?"). In response to the lack of consensus on the definition of impulsivity, the UPPS was developed. The scale includes four dimensions: (1) premeditation (i.e. tendency to think and reflect on the consequences of an act before engagement); (2) perseverance (i.e. an individual's ability to sustain attention on a task that may be boring or difficult); (3) sensation seeking (i.e. a tendency to enjoy and pursue activities that are exciting as well as an openness to trying new experiences); and (4) urgency (i.e. compromised ability to resist impulses that are driven by negative affect). The UPPS scale is one of the few self-report measures to attend to state as the urgency scale assesses impulsivity specific to negative affect. More recently, urgency has been expanded in a separate 5th dimension called positive urgency to acknowledge the role of positive affect in impulsive behavior.

The behavioral paradigms used to measure impulsivity and the various subconstructs resulting from these measures can be broadly divided into two categories: (1) those that measure impulsive choice or impulsive decision-making and (2) those that measure impulsive action or motoric impulsivity (as described above in the disinhibition section). The common principal among these first types of tasks (impulsive choice or decision-making) is that the subject has to choose between a safer strategy that will produce a greater final gain and a strategy of bigger immediate wins paired with possible penalties that could result in a smaller final reward. Multiple tasks exist to study impulsive decision-making, and one of the most common measures used is the delay discounting procedure. The concept of delay discounting offers a well-known behavioral operationalization of impulsivity based on the tenet that reward is discounted as a function of its delay, with those individuals evidencing the steepest discounting considered impulsive. Said more simply, impulsive individuals often prefer smaller immediate rewards over larger delayed rewards. Consequently, the highly impulsive person has the propensity to choose the course of action that maximizes immediate gains as they are either unlikely to take into account or give value to future gains, or to give up immediate satisfaction. In contrast, the self-controlled individual gives value to future gains and thus, more successfully delays gratification.

A common mean to assess delay discounting is by the delay discounting procedure, a paper/pencil-administered version of the original monetary-choice questionnaire. The questionnaire consists of a fixed set of 27 choices between smaller immediate rewards and larger delayed rewards. For example, participants are asked "Would you prefer $54 today or $55 in 117 days?" Participants are instructed to show preference by choosing an option, with the effect of delay resulting preference for smaller sooner amounts of money indicative of impulsivity. More recently, the experiential discounting

task (EDT) was created based on the adjusting amount procedure to assess delay discounting. In this task, participants make choices between real delayed and probabilistic amounts of money, delayed by 0, 7, 14, or 28 s. Choices are between a probabilistic (35% chance of receiving) delayed standard amount and an adjusting smaller amount that is immediate and certain. All choice consequences (delays, probabilities, and rewards) are experienced during the testing session, that is, while the participant is still making choices.

Both self-report and behavioral tasks have some limitations. Impulsivity self-report scales are subject to response bias. Specifically, with self-report measures, participants must recognize and report on their own behavioral tendencies in various contexts relative to other individuals, and these self-perceptions may not always accurately reflect their behavior. Self-reports require insight or cognitive ability to understand questions to provide an accurate report of behavior. In other words, an impulsive individual may not always be sufficiently reflective to perceive his or her own impulsivity. In contrast, performance on behavioral tasks is potentially more objective and thus less sensitive to biased self-perceptions. On the other hand, the behavioral tasks typically measure only one specific dimension of behavior (e.g. the value of delayed rewards or response inhibition), which may have limited generalization to broader behavioral contexts or to the multidimensional construct of impulsivity. Further, self-report generally is limited to a more global assessment and may miss current state effects, whereas the opposite is of concern with behavioral tasks in that they may be overly influenced by proximal factors (e.g. mood, setting).

Risk Taking

Risk taking can be described as engagement in socially defined problem behavior that may bring positive consequences to the individual, but may be considered undesirable by the norms of society and often includes the potential for harm or danger. Research on risk taking has encompassed a variety of behaviors including alcohol consumption, tobacco use, risky sexual activity, dangerous driving, interpersonal aggression, and delinquent behaviors. One line of research considers risk taking to result from poor decision-making processes. A well-validated strategy for assessing this perspective on risk taking is the Iowa gambling task (IGT). In the IGT, participants are instructed to pick cards from different decks. Some decks include cards signifying large wins, while other decks include cards with smaller wins. Over time the decks with the larger wins also begin to include an even larger loss (resulting in an overall loss and referred to as the disadvantageous decks) while the decks with the small wins start to also include a loss but one that is smaller than the win (resulting in an overall gain and referred to as the advantageous decks). Risky decision-making is measured by the number of selections from the disadvantageous decks.

As a compliment to this conceptualization of risk taking as poor decision-making, risk taking propensity is the consideration of risk behavior highlighting the balancing of positive and negative consequences associated with a particular risk. As such, risk taking propensity considers one's willingness to take risks along a continuum where some level of risk is beneficial and it is up to the individual to determine their own level of risk tolerance. The most common strategy for measuring risk taking propensity is the balloon analog risk task (BART), a computerized measure. In this task, the participant is presented with a display of a small balloon and asked to pump the balloon by clicking a button on the screen. With each click, the balloon inflates a small amount and actual money is added to the participant's temporary winnings. At any point, the participant has the option to press a button labeled "Collect $$$," which deposits the amount in temporary winnings to the bank (i.e. it can no longer be lost) and ends the trial, at which point a new trial begins. However, each balloon is programmed to pop somewhere between 1 and 128 pumps, with an average breakpoint of 64 pumps. If the participant fails to press "Collect $$$" before the balloon pops, all earnings for that balloon are lost and the next balloon is presented. Risk taking is defined as the average number of pumps on unpopped balloons, with higher scores indicating greater risk taking. The task was developed to provide a controlled setting in which to model risk taking in the natural environment, where risk taking up to a certain point leads to positive consequences, with further excessive risk taking leading to greater negative consequences that outweigh the positives. Risk taking propensity assessed in this way captures the appetitive processes underlying a behavioral tendency to take risks in response to cues for a potential reward with a probability for undesirable results, which has shown correlations with a variety of real world risk behaviors from adolescents through adults. There are several other frequently used tasks, but most capture similar dimensions to the IGT or BART and therefore are not discussed here.

It is important to consider the manner in which risk taking propensity overlaps with, and is distinct from, sensation seeking which is a construct also quite relevant for understanding the development of risk taking behavior. Sensation seeking refers to the tendency to seek out novel, varied, and highly stimulating experiences, and the willingness to take risks to attain them. Sensation seeking can be divided into four traits: thrill- and adventure-seeking, experience-seeking, disinhibition, and boredom susceptibility. Clearly, the drive to

seek out novel and exciting experiences fits well conceptually with a willingness to take risks more broadly. However, it is important to note clinical examples where a willingness to take certain risks such as not wearing a seatbelt have little to do with sensation seeking. Empirical data indicate only modest overlap between the constructs.

Link between Disinhibition, Impulsivity, and Risk Taking

The links between disinhibition, impulsivity, and risk taking are apparent from the overlap in their definitions. All involve diminished restraint or control and continued behavioral engagement despite adverse consequences. This link also relates to addiction which entails a compulsive pattern, characterized by impaired control over the use of the substance and or engagement in a behavior (e.g. gambling) and continued use/engagement in the face of negative consequences. Yet, performance on laboratory measures of impulsive choice (e.g. delay discounting) and disinhibition (e.g. stop and go task), and risk taking propensity (e.g. BART) are not necessarily related in humans. Although differences may result from measurement issues, they also may indicate different specific mechanisms underlying performance across these constructs. As mentioned briefly above, a recent study exploring the relationship between three widely used self-report measures and four laboratory-task measures found that the correlations among the various self-report measures were high, but self-reports were not correlated with behavioral task measures. Conducting a principal-components analysis on data from four behavioral tasks, two components emerged, labeled "impulsive decision-making" and "impulsive disinhibition." Impulsive decision-making included measures that involve participants making decisions about delayed versus immediate or probabilistic versus definite outcomes, namely delay discounting and the BART, while behavioral disinhibition included measures in which participants are expected to inhibit prepotent motor behavior that is the stop task and the go/no-go task (behavioral disinhibition task that measures withholding a response to avoid punishment). This study supports other recent findings indicating that self-report and behavioral tasks probably measure different constructs and suggest that even among the behavioral measures, different tasks measure different components of impulsive behavior.

RELATION TO ADDICTIVE BEHAVIORS

There is a large collection of research demonstrating a relationship between addictive behaviors and impulsivity and risk taking propensity. Drug users have consistently been found to have impaired inhibition compared with nonusers. Cocaine and methamphetamine users have displayed inhibitory deficits on the stop and go task compared with nonusers. Cocaine users and individuals with alcohol dependence have demonstrated decreased ability to inhibit prepotent responses on the go/no-go task compared with controls. Outside of illicit drug use, research has indicated that cigarette smokers show decreased inhibitory control compared with nonsmokers and that the number of packs smoked each day is positively related to inhibitory failures.

Delay discounting tasks have shown that alcohol, cocaine, opioid, methamphetamine, and cigarette users perform more impulsively relative to nonsubstance users. For example, in a study comparing opioid-dependent participants and nonuser participants, those who were opioid-dependent discounted delayed monetary rewards more frequently. Additionally, opioid-dependent individuals who share needles have been found to discount delayed rewards at a greater rate than their opioid-dependent counterparts who do not share needles.

In addition, substance users have performed more poorly on the IGT than control groups, as evidenced by lower overall scores in the task and their inability to shift game card selections to more advantageous decks. Research utilizing the BART suggests positive correlations with a range of addictive behaviors. For example, the BART has differentiated between nonsmokers and smokers, for which the latter displayed higher scores on the task, and has been related to a larger set of real world risk behaviors including several addictive behaviors among adolescents.

Characteristics of substance use, including frequency and quantity, have also been associated with impulsivity and risk taking propensity. For example, heavy drinkers have been shown to be more impulsive than light drinkers, and those who are more impulsive use greater quantities of cigarettes, alcohol, and marijuana. This finding has been replicated across diverse age groups. For example, among adolescents, those who discount delayed rewards at a greater rate have exhibited more illicit drug abuse. In a sample of college students, those that displayed stable, heavy-episodic drinking made fewer advantageous choices on the IGT compared with the low, binge drinking participants.

Animal models have also supported evidence linking impulsivity with addictive behaviors and greater substance use. Rats categorized as highly impulsive (assessed by a delay discounting procedure involving access to two response levers with varying delivery times and a food pellet dispenser) have been found to consume more of a 12% ethanol solution than rats with

lower impulsivity. In addition, rats that display higher impulsivity have self-administered cocaine more often and at a faster rate than those that were identified as lower impulsivity rats.

Impulsivity has also shown to have an effect on drug treatment outcomes, which are worse for those who have shown increased trait impulsivity. For example, impulsive adolescents (defined by discounting monetary rewards more on the EDT and committing more commission errors on a continuous performance task) have been found to be less likely to achieve smoking cessation in a 4-week treatment program compared with their less impulsive counterparts. Similarly, individuals that scored high on self-report measures of impulsivity were more likely to drop out of cocaine abuse treatment. They also remained in treatment for shorter lengths of time than those with lower impulsivity scores. Impulsive individuals also experience increased drug cravings during withdrawal, as well as a greater likelihood of relapse. For instance, smokers who scored high on impulsivity exhibited increased craving in response to cigarette cues and they relapsed more quickly than less impulsive smokers.

Prospective

Moving to understand causal relationships, there is a growing body of work suggesting that impulsivity and risk taking propensity predict addiction. For example, research with community samples has consistently shown that personality traits related to behavioral undercontrol including impulsivity, measured using various self-report inventories in early to late adolescence, predict the subsequent development of tobacco, alcohol, and illicit drug symptoms and disorders in early adulthood. However, it is not the case that these variables necessarily remain static over time, as more recent research suggests that risk taking propensity and sensation seeking change over time in adolescence and that change is related to change in real world substance use behaviors.

Animal studies have helped to clarify links between disinhibition and addictive behaviors. Researchers have investigated links between impulsivity and novelty seeking and the shift to compulsive cocaine use that characterizes human addiction. Researchers characterized rats as high or low responders based on their level of locomotor reactivity in a novel environment (thought to model sensation seeking) and as high or low in impulsivity based on their observed ability to delay performance in a five-choice serial reaction time task. They compared high and low responding rats and high and low impulsivity rats in their likelihood of initiating cocaine self-administration and in the transition to compulsive use (e.g. continued cocaine use despite negative consequences). They found that while high responders were more likely to initiate cocaine self-administration, high impulsivity predicted the propensity to transition from use to a compulsive pattern of use. These findings suggest that sensation seeking predicts cocaine consumption, while impulsivity predicts risk of developing an addiction.

Substance Use as a Causal Variable

Together, these data support the hypothesis that impulsivity and risk taking propensity predict addictive behaviors. However, the reverse relationship – impulsivity and risk taking are exacerbated by substance use – has also been supported within the literature. Theoretically, it has been posited that addiction results from an imbalance that occurs in the course of drug use in which the impulsive system comes to dominate the reflective system, leading to a decrement in the user's willpower to resist using drugs. Thus, addictive behaviors may drive maladaptive decision-making through pharmacologic interactions with neurophysiological mechanisms that evolved for the process of normal learning. Such decisions may at times be considered impulsive and risky. Substance use may initially be deliberate and involve significant planning, but with frequent repetition, these behaviors may transition to involving more rapid, unplanned, habitual actions. These habitual actions may be less sensitive to consequences. For addicted individuals, the decision to continue to use substances may reflect the habit system dominating the decision-making process. Thus, substance use may lead to decreased behavioral control. In addition, addictive behaviors may place an individual in situations in which impulsive, risk-prone actions are more likely, for example, being around other impulsive people or ending up in situations in which impulsive or risky decisions are more valuable given uncertainties in a chaotic environment.

In terms of empirical support, researchers have found that drug abstinence is associated with lower levels of impulsivity in substance users. In a study examining the effects of residential substance abuse treatment on behavioral control variables, risk taking propensity was found to significantly decrease from pre- to post-treatment. Furthermore, after a long-term period of abstinence from continued smoking, ex-smokers have been shown to discount delayed money rewards to the same extent as those who have never smoked and less than current smokers. These findings suggest that impulsivity and risk taking are heightened during periods of use, and return to nonsubstance using levels after cessation; however, other research suggests long-term impairment in impulsivity as a result of substance use. For example, researchers have compared the

performance of detoxified, alcohol-dependent patients with the performance of control participants in behavioral tasks measuring impulsivity. In comparison with the control participants, individuals with a history of substance abuse performed worse on the impulsivity measures. These results were interpreted by the researchers as suggesting that exposure to drugs is sufficient to cause long-lasting increases in impulsivity. Further research is needed to clarify these potentially conflicting findings.

To address the acute effects of drugs on impulsivity and risk taking behavior, researchers have begun to study the effects of these drugs administered in a lab setting, much of which has focused on alcohol. For example, there is a body of work examining the effects of acute alcohol administration on specific state-dependent measures of impulsivity showing that low, moderate, and high doses of alcohol increase impulsive performance among moderate alcohol drinkers recruited from the community. Similarly, a moderate dose of alcohol has been found to increase response inhibition failures measured among college students. Further, among college students defined as binge drinkers, alcohol has increased response inhibition failures relative to nonbinge drinkers. On the other hand, studies examining the relationship and effects of alcohol on delay discounting measures have shown mixed results. After consumption of either moderate or large alcohol doses, increased impulsive responding on delay discounting measures (i.e. increased numbers of shorter-sooner reward choices) was found in a sample of alcohol abusers, but among healthy adults, there was no effect on impulsive responding. Data are less clear with risk taking measures. Although more research is needed to draw firm conclusions, studies conducted largely have not shown a clear and reliable effect of alcohol on risk taking. Interestingly, although a smaller body of research, studies examining stimulant drugs actually seem to indicate a decrease in impulsivity and risk taking. Although counterintuitive in some ways, this effect mirrors the effect of stimulant drugs on impulsive behavior in clinical conditions such as ADHD and more research is needed to fully understand the nature of these relationships.

Common Underlying Cause

To summarize, research has demonstrated a relationship between impulsivity and risk taking propensity and addiction; however, the exact nature of the association warrants clarification. There is available evidence that these variables may act as a precursor to the onset of addiction; yet, addictive behaviors may also lead to increased risk taking and impulsivity. A third possibility is that the association between disinhibition and addiction may be explained by common etiologies, an explanation we now turn to in more detail.

Behavioral Genetics

Such common etiologies may present at the genetic level. It is commonly thought that addiction has a genetic component. Family and twin epidemiologic studies have estimated that genetic contributions can account for up to 60% of the variance in the risk for addictions. Relatives of probands with substance use disorders have been found to be at an eightfold increased rate for these disorders. Yet, it has been difficult to replicate reports of specific genes for substance use disorders as genetic association studies have been characterized by inconsistent results which have raised interest in more complex approaches including larger scale genome-wide association studies (GWAS) and efforts to consider gene by environment interactions. Additionally, researchers have become increasingly interested in the variables that serve as the intermediary step between genes and substance use disorders; referred to as endophenotypes, risk taking propensity and impulsivity serve as two especially promising examples.

Research has shown addiction to be highly correlated with other psychological disorders in which impulse control is implicated. Compelling evidence suggests that problems such as substance use disorders, attention deficit hyperactivity disorder, conduct disorder, and antisocial personality disorder, among others, should be classified together under the rubric of externalizing disorders. This argument is based on evidence indicating that specific externalizing syndromes are often comorbid. For example, a national survey consisting of participants between 15 and 54 years old showed that alcohol dependence, drug dependence, and antisocial personality disorder formed a single externalizing factor, which could be distinguished from a separate factor that represented internalizing disorders (e.g. major depression, generalized anxiety disorders). Among externalizing disorders, impulsivity serves as a common link.

Additional support for classifying these disorders under one rubric comes from genetic correlations that have been observed among externalizing disorders, including alcohol abuse/dependence, drug abuse/dependence, and the child and adult aspects of antisocial personality disorder. Research has demonstrated that the genetic influences on the personality trait of behavioral undercontrol accounted for the majority (90%) of the genetic risk factors common to conduct disorder and alcohol dependence. Other research has presented significant genetic correlations among the personality traits of social deviance and excitement seeking, and alcohol consumption and problems, in young adult Finnish twins. Similarly, significant genetic correlations linking alcohol misuse and a variety of

personality scales indexing antisocial tendencies has been found. Finally, in the Virginia Twin Cohort, a common factor encompassing externalizing disorders such as antisocial personality disorder and conduct disorder accounted for 71% of the genetic liability to alcoholism and 67% of the inheritance of vulnerability to illicit drug abuse and dependence. While these findings point to main effects of genes in relation to externalizing psychopathology, research also indicates the likelihood of environmental experience moderating the effects of specific genes.

Neurotransmitter Systems

A number of systems are implicated in the connection between impulse control and addictions, including dopamine, GABA and glutamate, serotonin, and MAO. For example, genetic studies have linked several dopamine-related genes to impulsivity and addiction, including genes encoding the DA D4 receptor (*DRD4*) and dopamine transporter (*SLC6A3*). Additionally, the *D2A1* allele of the D2 receptor has been implicated in drug abuse smoking and is significantly more common in subjects with problem gambling compared with controls. Additionally, alterations in dopaminergic pathways have been proposed as underlying the seeking of rewards that trigger activation of the dopaminergic system, producing feelings of pleasure.

γ-aminobutyric acid (GABA), which is the main inhibitory neurotransmitter in the brain, has also been implicated in impulse control disorders and addiction. There is evidence of anatomic and functional connectivity between GABA and dopaminergic systems as well as increasing support for effects of modulation of GABAergic systems on substance use disorders. For example, tiagabine, a GABA reuptake inhibitor used primarily to treat seizures, has shown preliminary efficacy in cocaine addiction, and in a case report, was shown to help with control of impulsive aggression. Further, glutamate, an excitatory neurotransmitter and precursor of GABA, has also been associated with addictions. In preclinical studies, levels of glutamate within the nucleus accumbens mediate reward-seeking behavior. Taken together, these data suggest possible roles for glutamatergic and GABAergic systems in addictions.

In addition, a role for serotonin (5-HT) has also been supported in both impulsivity and addictions. In animal models, forebrain 5-HT depletion has been shown to lead to impulsive choice, while the indirect 5-HT agonist fenfluramine decreases such behavior. Nonselective 5-HT antagonists have been shown to promote self-controlled choice. Low levels of the 5-HT metabolite 5-hydroxyindolacetic acid (5-HIAA) have been found in individuals with impulsive characteristics and early-onset alcoholism. Low levels of cerebrospinal fluid 5-HIAA have also been associated with risk taking behaviors in primates (e.g. monkeys taking longer leaps in the jungle, violent criminals, and impulsive suicide committers). In general, low levels of brain 5-HT activity have been associated with impulsive behaviors, such as outward and self-directed violence and aggression, suicide, fire starting, and pathological gambling. Taken together, multiple lines of evidence support a role for 5-HT in mediating impulsivity.

Several studies have also demonstrated a strong link between low platelet monoamine oxidase (MAO) activity and drug use, alcohol dependence, antisocial behaviors, suicide, and gambling. In addition, low levels of MAO have also correlated with self-reported impulsivity and novelty seeking. In a sample of 557 young adult males aged 21–26, low MAO activity was associated with low impulsive control as indicated by high scores on measures of antisocial tendencies, higher rates of cigarette smoking, and increased rates of drug use. However, low MAO activity was not related to either alcohol use or abuse.

Developmental and Neurodevelopmental Factors

Developmental research suggests that, in general, impulse control continues to mature over the course of adolescence and early adulthood. For example, in a sample of individuals ranging in age from 7 to 29, a significant negative correlation between chronological age and impulsivity (using the Connors impulsivity scale) was found. Similarly, a significant decline in impulsivity from ages 14–16 to 20–22 on the EIS and the BIS has been observed. This was recently replicated in a large, socioeconomically and ethnically diverse sample of individuals between the ages of 10 and 30. Specifically, using both self-report and behavioral measures, impulsivity followed a linear pattern, declining steadily from age 10 on (as a note of caution these are cross-sectional findings). Delay discounting was also examined within this sample; results demonstrated that younger adolescents had a greater willingness to accept a smaller reward delivered sooner than a larger one that is delayed than did individuals aged 16 and older. Similar delay discounting findings have been found with other between-group cross-sectional comparisons of young adolescents (age 12), young adults (average age 20 years old), and older adults (average age 70 years old). These studies demonstrate that young adolescents discount monetary rewards more steeply than young adults, who discount at a faster rate than older adults. Together these findings suggest a decline in impulsivity from childhood through adolescence and into adulthood.

These findings parallel evidence showing structural and functional maturation over the course of adolescence in brain regions that subserve impulse control and other aspects of self-regulation. The development

of the prefrontal cortex (PFC) is believed to play an important role in the maturation of higher cognitive abilities such as decision-making and cognitive control, processes inherently related to impulse control. A number of paradigms, including response inhibition tasks, have been used, in conjunction with functional magnetic resonance imaging (fMRI), to assess the neurobiological basis of these abilities. Together, these studies show that children recruit distinct but often larger, more diffuse prefrontal regions when performing these tasks than do adults. The pattern of activity becomes more fine-tuned with age, with regions not correlated with task performance diminishing in activity with age.

In considering these findings in relation to observed increased risk taking behavior, including addictive behaviors, in adolescence, there has been a recent focus within neuroimaging studies to examine reward-related processing, focusing primarily on the region of the accumbens (a portion of the basal ganglia involved in predicting reward) in conjunction with top-down control regions (i.e. PFC). For example, a recent study examined behavioral and neural responses to reward manipulations across development. Results suggested enhanced accumbens activity in anticipation of rewards during adolescence relative to children and adults. Further, both children and adolescents showed a less mature response in prefrontal control regions than adults. The authors concluded that these findings suggest different developmental trajectories for these regions which may underlie the enhancement in accumbens activity and in turn relate to the increased risky behaviors observed during this period of development.

Thus, two core neurobiological networks seem to underlie human decision-making and are important in the emergence of risky behavior, including addictive behavior, observed in adolescence. The first, a cognitive control system, which consists of prefrontal and parietal regions, as well as the anterior cingulate, facilitates executive functioning. The second, affective system includes regions which are important to processing reward and social and emotional salience, including but not limited to the amygdala, ventral striatum, orbitofrontal cortex, medial PFC, and the superior temporal sulcus. Empirical evidence supports the notion that it is the affective network which dominates during adolescence, with heightened reward sensitivity throughout this developmental period.

The emergence of this affective network coincides with the onset of puberty, while the cognitive control network develops across a longer developmental trajectory, further supporting the dominance of the affective system in adolescent decision-making and increasing risky behavior. This dissociation between an affective and cognitive system has also been conceptualized as dissociation between an activational system and an inhibitory system, with delayed development of the inhibitory system and the dominance of the activational system being responsible for the onset of risky behaviors. In sum, the staggered development of these two brain regions is thought to result in heightened reward drive and cognitive disinhibition leading to increased engagement in addictive behavior.

In adolescents with conduct problems and adults with alcohol use disorders, reports from electrophysiological studies have found an association between inhibitory control impairment and PFC dysfunction. In an extensive review of imaging studies, it was reported that neuroimaging techniques have consistently documented differential activation in PFC in individuals with substance use disorders compared with controls when performing tasks of cognitive inhibition. Findings have also documented smaller PFC volumes across different studies in individuals with substance use disorders relative to controls. The PFC has also been associated with tolerance of delayed rewards.

CONCLUSION

Impulsivity (including disinhibition) and risk taking propensity represent complex, multifaceted constructs with relevance to addictions. Researchers are refining the definitions and assessment of different subtypes of impulsive behavior and relating these to the causes and consequences of addiction-related behaviors and disorders. The relationships may be best described as reciprocal and bidirectional, where these constructs are viewed as determinants and consequences of addictive behaviors, as well as concurrent factors emanating from similar causal mechanisms with particular relevance of biological, individual difference, and developmental factors. As research leads to a better understanding of these complex relationships, there will be much needed advancements in how we assess and treat addictive behaviors, which have the potential to contribute to advancements in public health.

SEE ALSO

Alcohol Use Disorders, Heroin Addiction, Cocaine Addiction, Marijuana Use and Abuse, Methamphetamine Addiction, Ecstasy/MDMA, Tobacco, Gambling, The Biopsychosocial Model of Addiction, Personality and Addiction Processes, Developmental Risk Taking and the Natural History of Alcohol and Drug Use Among Youth

List of Abbreviations

5-HIAA	5-hydroxyindolacetic acid
5-HT	serotonin
BART	balloon analogue risk task
BIS	Barratt impulsiveness scale
DRD4	dopamine receptor D4
EDT	experiential discounting task
EIS	Eysenck impulsiveness scale
fMRI	functional magnetic resonance imaging
GABA	γ-aminobutyric acid
GWAS	genome-wide association studies
IGT	Iowa gambling task
MAO	monoamine oxidase
PFC	prefrontal cortex
RTs	reaction times
SLC6A3	dopamine transporter
UPPS	UPPS impulsive behavior scale

Glossary

Accumbens a collection of neurons within the striatum. It is thought to play an important role in reward, the experience of pleasure, laughter, addiction, aggression, fear, and the placebo effect.

Anterior cingulate cortex a part of the brain located in the frontal part of the cingulate cortex. It plays an important role on automatic and cognitive functions, including reward anticipation and decision-making.

Appetitive processes pleasant or wanted stimulus that a person will naturally try to approach.

Dopamine neurotransmitter that has been associated with cognition, motivation, learning, punishment, and rewards.

Functional magnetic resonance imaging Neuroimaging technique that focuses on blood flow in the brain to detect areas of activity.

γ-Aminobutyric acid main inhibitory neurotransmitter in the brain, which has been implicated in impulse control disorders and addiction.

Glutamate an excitatory neurotransmitter and precursor of GABA, which has also been associated with addictions.

Monoamine oxidase an enzyme that catalyzes the oxidation of monoamines. Low platelet MAO activity has been associated with drug use, alcohol dependence, antisocial behaviors, suicide, and gambling.

Orbitofrontal cortex prefrontal cortex brain region in the frontal lobes. It is involved in the cognitive processing of decision-making.

Prefrontal cortex anterior part of the frontal lobes of the brain responsible for executive functioning.

Reward drive argued to reflect individual differences in sensitivity to incentive motivation and engagement in appetitive behavior on detection of reward cues.

Serotonin a neurotransmitter involved in transmitting nerve signals between nerve cells and that causes blood vessels to narrow. It is thought to play an important part in impulsivity and addictions.

Urgency impulsivity accompanying negative affect.

Further Reading

Brewer, J.A., Potenza, M.N., 2008. The neurobiology and genetics of impulse control disorders: relationships to drug addictions. Biochemical Pharmacology 75, 63–75.

Crews, F.T., Boettiger, C.A., 2009. Impulsivity, frontal lobes and risk for addiction. Pharmacology, Biochemistry and Behavior 93, 237–247.

Galvan, A., Hare, T.A., Parra, C.E., et al., 2006. Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. Journal of Neuroscience 26, 6885–6892.

Iacono, W.G., Malone, S.M., McGue, M., 2008. Behavioral disinhibition and the development of early-onset addiction: common and specific influences. Annual Review of Clinical Psychology 4, 325–348.

Kreek, M.J., Nielsen, D.A., Butelman, E.R., LaForge, K.S., 2005. Genetic influences on impulsivity, risk taking, stress responsivity and vulnerability to drug abuse and addiction. Nature Neuroscience 8, 1450–1457.

Krueger, R.F., Hicks, B.M., Patrick, C.J., et al., 2002. Etiologic connections among substance dependence, antisocial behavior and personality: modeling the externalizing spectrum. Journal of Abnormal Psychology 111, 411–424.

Leeman, R.F., Grant, J.E., Potenza, M.N., 2009. Behavioral and neurological foundations for the moral and legal implications of intoxication, addictive behaviors and disinhibition. Behavioral Sciences & the Law 27, 237–259.

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). Journal of Experimental Psychology: Applied 8, 75–84.

Moeller, G.F., Dougherty, D.M., 2002. Impulsivity and substance abuse: what's the connection? Addictive Diseases Treatment 1, 3–10.

Perry, J.L., Carroll, M.E., 2008. The role of impulsive behavior in drug abuse. Psychopharmacology 200, 1–26.

Reynolds, B., Ortengren, A., Richards, J.B., de Wit, H., 2006. Dimensions of impulsive behavior: personality and behavioral measures. Personality and Individual Differences 40, 305–315.

Richards, J.B., Zhang, L., Mitchell, S., de Wit, H., 1999. Delay and probability discounting in a model of impulsive behavior: effect of alcohol. Journal of the Experimental Analysis of Behavior 71, 121–143.

Steinberg, L., Albert, D., Cauffman, E., et al., 2008. Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: evidence for a dual systems model. Developmental Psychology 44, 1764–1777.

Volkow, N.D., Fowler, J.S., Wang, G.J., Baler, R., Telang, F., 2009. Imaging dopamine's role in drug abuse and addiction. Neuropharmacology 56, 3–8.

Whiteside, S.P., Lynam, D.R., 2001. The five factor model and impulsivity: using a structural model of personality to understand impulsivity. Personality and Individual Differences 30, 669–689.

CHAPTER

22

Emotions and Addictive Processes

Jon D. Kassel, Jennifer C. Veilleux, Adrienne J. Heinz, Ashley R. Braun, Megan Conrad

University of Illinois at Chicago, Chicago, IL, USA

OUTLINE

INTRODUCTION: THE COMPLEX RELATIONSHIP BETWEEN EMOTIONS AND ADDICTIVE PROCESSES

Individuals take drugs for a variety of reasons, including peer pressure, genetic transmission of vulnerability to drug effects, parental modeling, and celebratory (positive reinforcing) rituals, to name but a few. The vast majority of people who use drugs (be they legal or illicit) do not ultimately transition to a state of drug addiction or dependence. Drug use more typically serves as an experimental rite of passage, or an occasional behavior marked by social engagement.

However, and as we discuss in more detail below, those unfortunate individuals who progress to the deleterious stage of drug addiction are the focus of this chapter. But why examine the relationship between drug addiction and emotion? What is to be gained by such an inquiry? In response to these questions, we argue that of all of the reasons (or attributions) provided by people who abuse drugs, the relationship between addiction and emotion may ultimately prove the most important and salient. Hence, if one were to simply ask abusers and addicts why they use drugs, the modal response would typically reflect aspects of the extent to which drugs make them feel something, be it the

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00022-X

removal of unpleasant affect (negative reinforcement), or the heightening of positive affect (positive reinforcement). Such associations between feeling states and addictive behaviors have been well studied in the empirical literature and have important implications for prevention and intervention efforts.

This chapter begins with an overview of several important terms: emotion and addiction. We then examine the comorbidity of drug abuse problems and disorders of affect (e.g. depression, anxiety) and consider several influential theoretical models positing critical associations between drug use and emotional states. Next, we provide a brief review of the evidence on whether drugs genuinely do render influence on emotional states. The chapter concludes with a summary discussion and future research directions.

CLARIFICATION OF TERMS: ADDICTION AND EMOTION

Emotion

Defining the constructs of emotion and addiction is essential to understanding the multidirectional relationship between emotion and addictive processes. Although most frequently conceptualized as subjective feeling states, emotions are multifaceted experiences that also include facial expressions, changes in autonomic arousal, cognitive appraisal processes, and behavioral sequelae. Generally, emotional processes are differentiated by the terms affect, mood, and emotion. Affect is the most general term, often used as a superordinate category to describe any emotional process, including basic evaluation of environmental stimuli as positive or negative. Moods are longer-lasting states of limited intensity that do not typically have a discernible causal object. Emotions are intense but short-lived acute states that are directly prompted by specific environmental experiences and/or objects. The basic emotional states (anger, fear, sadness, disgust, surprise, and happiness) are thought to be associated with particular patterns of neural responding that cut across cultures and can be observed in numerous species.

The terms emotion, mood, and affect describe relatively acute states, to be differentiated from affective traits and disorders of affect. Affective personality traits are longstanding, typically unconscious predispositions that influence the likelihood of experiencing an emotional response. For example, individuals high in neuroticism are more likely to experience negative emotional responses to stress and have more difficulty regulating their emotions. They are also more likely to experience disorders of affect such as mood disorders (depression, bipolar disorder) and anxiety disorders. Disorders of affect reflect pervasive alterations of mood that confer deleterious effects on thoughts, emotions, and behaviors, and frequently require psychotropic medication or psychotherapeutic intervention.

Addiction

Drug and alcohol use are not synonymous with addiction, although use is a necessary pre-requisite for development of addictive processes. Addiction reflects a maladaptive pattern of repeated use that results in significant harmful consequences. The American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition, Text Revision)* (DSM-IV-TR) has identified 11 classes of substances and a series of diagnoses associated with adverse effects of substance use including abuse, dependence, intoxication, withdrawal, and delirium. The term addiction is most closely tied to substance dependence disorders, characterized by at least three of the following symptoms: physical tolerance, withdrawal, greater use than intended, inability to reduce or control use, devotion of time and resources to obtaining the drug, and reduced social or occupational functioning due to drug use and continued use despite knowledge of the harmful consequences.

The terms addiction and dependence, particularly as differentiated from the use and abuse of substances, reflect underlying changes in brain functioning due to repeated substance use. Physiologically, drugs and alcohol influence brain systems and neurotransmitters associated with emotion (serotonin, dopamine, norepinephrine, etc.), with prolonged use appearing to exert long-term effects on the brain's emotional circuitry. Individuals addicted to drugs or alcohol may, therefore, have a biological need for the drug to maintain a consistent level of functioning, including in the realm of emotional processing.

Comorbidity

The relationship between addiction and emotion is likely bidirectional, reflected in the acute influences of heightened emotion on self-administration of substances, and the complimentary effect of substance use on immediate emotional responding. At a broader level of analysis, epidemiological studies have established a clear link between addictive disorders and disorders of affect. For example, rates of anxiety and depression are considerably higher in samples of substance-abusing individuals compared with the general population. Moreover, affective distress appears to increase in conjunction with the severity of substance issues. The

prevalence of anxiety and mood disorders is higher among those with substance dependence than substance abuse; individuals with multiple affective disorders also manifest more drug problems than people with a single mood or anxiety diagnosis. However, evidence is mixed as to the directional nature of these processes, such that affective disorders may confer vulnerability to substance use, and/or addictive processes may result in disorders of affect.

Part of the difficulty in elucidating causal patterns may be due to differences in symptomatic presentations across the specific disorders of affect, and/or differences among classes of drugs. For example, comorbidity rates of substance use disorders with mood disorders are typically higher than for anxiety disorders. Moreover, comorbidity rates differ by specific disorder; drug dependence rates are generally higher for depression than bipolar disorder, and higher for social phobia, simple phobia, and generalized anxiety disorder compared with panic disorder. Similar differences in comorbidity occur by drug classes, such that rates of affective disorders are higher for individuals addicted to drugs such as inhalants, tranquilizers, and sedatives compared with more commonly used drugs such as alcohol and marijuana. These nuances are important because they speak to different mechanisms that may drive the association between drug use and affective distress, and suggest that future research will need to solidify patterns of comorbidity as well as underlying processes that may account for the high rates of co-occurrence.

The relationship between disorders of affect and addictive behavior is also influenced by other moderating factors. For example, female addicts experience a higher degree of affective distress than male addicts, whereas men who abuse substances manifest higher rates of externalizing problems (i.e. antisocial personality disorder) than women. Comorbidity rates also vary by country and among sociodemographic factors within a given country (e.g. ethnicity, socioeconomic status). Moreover, factors such as childhood adversities, temperament, and peer and family processes all alter vulnerabilities to both substance use initiation and affective distress.

Delineating the casual direction of addiction and affective disorders suggests that the two are theoretically independent. However, several of the symptoms overlap, making clear distinctions between the disorders difficult to discern. For example, irritability and difficulty concentrating are frequent symptoms of depression, and are hallmark features of many drug withdrawal syndromes. To accurately assess factors associated with comorbidity, researchers must differentiate substance-induced psychiatric conditions that are direct sequelae of substance use or withdrawal from independent disorders that occur or persist more than 4 months after amelioration of withdrawal. Other methodological decisions have a bearing on the interpretation of epidemiological findings, including the choice of clinical or general population level samples, statistical control of sociodemographic factors such as race or income, and consideration of the high degree of comorbidity between anxiety and depression.

The literature on comorbidity between disorders of affect and substance use is definitive in one regard: there is clearly a link between emotion and addiction. It appears that affective distress begets problems with substance use, and substance use disorders lead to problems of affect. When viewed through a biopsychosocial lens, these findings are not surprising. At a neurochemical level, affective disorders and recurrent substance use cause altered brain chemistry, typically in brain regions associated with emotional processing. At a psychological and sociological level, addicts regularly encounter relationship conflicts, financial woes, and engage in other risky behaviors that may lead to prolonged emotional pain. Correspondingly, individuals suffering from affective disorders may choose to use drugs or alcohol as a way of self-medicating emotional distress. The high comorbidity rates between these classes of psychiatric conditions suggest the need for further research across multiple levels of analysis and methodologies (e.g. epidemiological, developmental, within-person, laboratory based, etc.) to further elucidate the bidirectional relationship between affective distress and addiction.

THEORETICAL FRAMEWORKS FOR THE ROLE OF EMOTION IN ADDICTIVE PROCESSES

Negative Reinforcement Models

Steeped in the traditions of Skinner's operant conditioning, negative reinforcement models of substance use and dependence posit that on learning that drug administration is followed by a reduction in distress, an individual will be more likely to use substances when distressed in the future. Hence, a vicious cycle is initiated and a seemingly intractable addiction emerges. This conceptualization is also represented within a web of related hypotheses (i.e. self-medication, stress coping, tension reduction) that collectively espouse that motivation to use substances is driven by promise of a key emotional pay-off: alleviation of negative affect (NA).

Withdrawal-Relief Model

The experience of drug withdrawal is typically characterized by NA symptoms that include depression,

anxiety, and irritability among other unpleasant mood states. The withdrawal-relief model, one of the oldest and most enduring proposed explanations for drug addiction, suggests that following repeated substance use, individuals begin to use substances for the purpose of avoiding withdrawal symptoms, and not to ameliorate organic NA. Although a compelling framework for understanding the later stages of drug dependence, the withdrawal-relief model does not explain why relapse can occur well after withdrawal symptoms have subsided. Moreover, it fails to adequately explain how patterns of use are established well before the onset of physical dependence and withdrawal (e.g. experimental and initiation phases). Most individuals in the early developmental stages report using substances to enhance positive mood (i.e. positive reinforcement).

Classic Conditioning and Cue Pairing

Principles of associative learning can be applied to account for some of the limitations associated with models that view negative reinforcement as the sole motivational core of substance use. For instance, exposure to drug cues previously associated with withdrawal symptoms can elicit a conditioned withdrawal response and subsequently evoke craving. More importantly, affective distress can signal an individual to use drugs even in the absence of withdrawal symptoms. This is because with repeated use, NA can eventually function as a warning flag that signals to the user that substances are typically consumed in this context.

During active use, the acute effects of a substance serve as an unconditioned stimulus (UCS). Initially, when a substance (UCS) is introduced to the system, regulatory processes attempt to restore homeostasis (i.e. unconditioned response (UCR)). After repeated pairings, the strength between the UCS and the UCR increases and the link between them becomes cognitively, emotionally, and physiologically entrenched. Eventually, the substance itself becomes a conditioned stimulus and the defensive regulatory response becomes the conditioned response. Once these associations become fully conditioned, exposure to the substance – or even substance-related cues – can elicit a compensatory response even in the absence of the individual actually having to use the substance. This response may be experienced as withdrawal, craving, and increased NA, all of which promote and heighten the likelihood of future substance use.

Reformulated Model of Negative Reinforcement (RMNR)

As mentioned earlier, it is generally accepted that as substance levels in the system drop and a withdrawal state emerges, the user begins to experience NA. Accordingly, these NA cues are believed to signal urges and promote substance use. The RMNR offers several important nuances to this conceptualization and asserts that avoidance of the affective components of withdrawal is the primary motivating force behind substance use. More specifically, the model proposes that as experience with substance use mounts, the user becomes more attuned and sensitive to the intereoceptive cues (i.e. internal sensations) associated with withdrawal symptoms. This withdrawal-associated internal state eventually becomes a potent conditioned stimulus that elicits NA. In the early stages of withdrawal, NA cues are subtle and fail to be perceived on a conscious level. Instead, the user's preconscious detection of the cues start to bias responding in a way that increases the probability that substance use, a previously reinforced response, will occur. In other words, once signs of withdrawal are preconsciously detected, the autopilot switch is flipped on in anticipation of affective distress. Subsequently, the individual's motivational processes are swayed toward substance use ostensibly to relieve or avoid the NA caused by declining levels of drug in the system.

These processes transpire outside of the user's awareness in an automatized fashion, that is, until a notable environmental stressor is experienced (e.g. interpersonal conflict) or access to a substance becomes challenged (e.g. smoking and air travel). At this point, the user crosses an affective threshold and enters into conscious awareness of the NA. Once the presence of NA is noted, urgency intensifies and the user's ability to process information is compromised. Attention and response selection are biased toward hot information – information regarding responses that have decreased NA in the past (i.e. substance use). Consequently, information about alternatives to procuring affective relief without substances (i.e. cold information) is deemphasized (e.g. going to a 12-step meeting). In addition, the influence of cognitive control (e.g. attention to goal-relevant information, action monitoring) on decision making is significantly diminished, resulting in increased likelihood of substance use.

The RMNR clearly takes a more ecological approach to describing the role of emotion in substance use in that it considers both endogenous (e.g. withdrawal symptoms) and exogenous factors (e.g. environmental stressors). It is thought that their inclusion can rectify some of the limitations noted in other negative reinforcement models, namely, that substance use can be induced by nonpharmacological affective distress – a stance that can explain why relapse may occur well after withdrawal-related aversive states have abated.

Positive Reinforcement Models

Traditional positive reinforcement models of drug motivation assert that individuals use drugs primarily

because they confer emotional benefit (rather than decreasing affective distress or relieving withdrawal symptoms). A wealth of neurobiological evidence shows that substance use induces a hedonic effect by increasing the levels of dopamine in the brain's reward center. Hence, drug users frequently describe their subjective experience as highly pleasurable, and this supports the idea that substance use is related to the neural basis of positive affect. The reward learning hypothesis suggests that dopamine plays a key role in teaching or stamping in that certain responses are followed by reward. That is, individuals become conditioned or programmed to repeat behaviors that result in reward. Once a substance has been associated with reward, the drug itself, along with associated cues, functions to activate the reward system. Such a framework may be particularly helpful in describing the processes governing early drug use (typically in adolescence), where positive reinforcement typically drives consumption.

Problems with Purely Positive or Negative Reinforcement Models of Addiction

Several inconsistencies have been highlighted in the literature regarding the models of drug motivation described above. First, withdrawal and craving do not reliably predict drug use. Furthermore, relapse is also known to occur in the context of positive affect (i.e. celebration). Second, many individuals express that they continue to use substances even when they no longer derive any affective benefit. Some substance users go so far as to express a genuine hatred of their drug use despite engaging in such use every day. Such observations suggest that negative reinforcement or positive reinforcement alone may not be the only mechanisms governing substance use. Other reinforcement models of substance use hint toward an interplay of positive and negative reinforcement processes in combination with key neurological and physiological processes. Such blended perspectives actually function to address many of the problems noted with models that cite only negative or positive reinforcement as the primary motives for substance use.

Blended Models

Opponent Process Model

The opponent process model of addiction is based on the premise that substance use disturbs individuals' natural balance (i.e. homeostasis). Initially this disturbance is in a positive and hedonic direction (the *a* process) although eventually the *b* process, the opponent process, is enacted to restore homeostasis. The difference in magnitude between the *a* and *b* processes determines the user's affective response to the substance. For instance, if the *a* process is larger in magnitude than the *b* process, the user is believed to experience a pleasurable state. Conversely, if the *b* process is larger, it is thought the user will experience a dysphoric affective state. In the initial phases of substance use, the appetitive a process appears more robust and ostensibly reinforces the individual to continue using. However, over time with repeated use and growing tolerance, the *b* process begins to dominate the *a* process. This shift results in a negative net effect where disrupted homeostasis manifests as withdrawal symptoms. Accordingly, abstinence is the only means by which the *a* and *b* processes can be restored to baseline levels.

Allostatic Reward Model

Derived from animal research, this model claims that addictive processes cannot be fully explained by homeostatic functions because addiction is inherently an allostatic state. That is, chronic drug use moves the homeostatic reward set point away from its naturally calibrated position. More specifically, in the development of addiction, functional changes are induced in the brain's reward system and other neuroadaptive mechanisms. Although compensatory processes are at work (e.g. the opponent process), they eventually fail to bring the user back to a normal homeostatic range. The individual then begins to derive a disproportionate amount of reward from substance use relative to other rewards such as food, sex, etc. An allostatic state is achieved when the brain circuitry becomes so dysregulated that it can no longer maintain stable reward functioning and the reward set point remains chronically off. When in this state, hormonal stress responses are triggered and activation of the cortico-striatal-thalamic loop among other brain areas promotes compulsive drug seeking and eventual loss of control. Overall, the model offers a viable explanation for the frequently observed downward spiral of distress experienced by addicts. The allostatic reward model also goes some way toward explaining why individuals are neurologically vulnerable to relapse even after extended periods of abstinence.

Incentive Sensitization/Salience Model

A common observation in clinical settings is that substance users continue engaging in use even though they report not liking it. For instance, some cigarette smokers express a deep hatred of smoking or report trivial affective benefits, yet continue to smoke regularly and without serious intention of quitting. In addition, users can recall catching themselves starting to use after a long period of abstinence without even consciously wanting to do so. Such marked discrepancies between

expressed motivational intent and observed objective behavior are puzzling and challenge the tenets of traditional motivational models. Ultimately however, this anecdotal evidence has served as a foundation for researchers to examine whether differences between liking a substance and wanting a substance could illuminate the pathways to abuse, dependence, and even relapse.

In keeping with the notion that persistent and complex changes occur in the brain over the course of addiction, the incentive sensitization theory posits that compulsive substance use is caused by a progressive hypersensitization of neural mechanisms (involved in reward) to the pleasurable effects of a substance (e.g. liking). In the initial phases of drug use, this means that motivational significance, or reward salience, of the desirable effects of the substance begins to shift such that increased reward is derived from each subsequent use (i.e. sensitization). As hypersensitivity to the incentives associated with the substance develops, attentional processes become strongly biased toward the substance as well as substance-related stimuli. This shift, coupled with other known substance-related impairments in executive functioning, induces a compulsive wanting of the drug that is thought to constitute the pathological core of addiction. In effect, the system becomes conditioned to want the substance in a way so powerful that it can occur outside the realm of conscious awareness and even after long periods of abstinence. The critical piece here is that neural mechanisms involved in the pleasurable (liking) effects of substance use do not become sensitized nearly to the same extent as wanting. Burgeoning evidence supports the presence of neural mechanisms underlying the subjective, pleasurable (liking) effects of a substance that are distinct from mechanisms associated with reinforcing (wanting) effects.

Cognitive Models

Although cognitive theories of substance use are well developed, their intersection with emotional processes is understood to a somewhat lesser extent. More commonly, specific cognitive mechanisms are identified as mediators in the relationship between emotion and substance use. Such an approach honors the inherent complexities that characterize the role of emotion in substance use, and may help to explain why emotional response to the same substance-related stimuli or internal states can vary so greatly within the same individual over time and across contexts. Moreover, cognitive models appear to address some of the unexplained variance in situations where the effects of emotion fail to demonstrate predictable influences on substance use.

Expectancy Theory

Data from decades of research with placebo designs undoubtedly attest to the notion that what one expects to happen tends to strongly influence what actually happens. Accordingly, several cognitive mechanisms have been proffered to explain why expectations so strongly dictate experience. In the case of drug outcome expectancies, individuals formulate beliefs about the affective consequences of using a substance (e.g. feeling relaxed after a cocktail). These expectancies are acquired through social learning and media messages, and are shaped by repeated experiences of positive and negative reinforcement with a substance. As the user becomes more experienced, connections between schematic representations (i.e. cognitive framework for organizing information) for substances and affective expectations strengthen in the brain. These affectively laden expectancies lie at implicit as well as explicit levels of awareness. When activated in memory, they are thought to govern substance-seeking behavior in a seemingly automatic fashion.

Attention Allocation Model

Based on the observation that, both within and across individuals, the effects of alcohol on a host of cognitive, behavioral, and emotional outcomes seem inconsistent, Claude Steele and colleagues formulated the influential attention allocation model. Simply put, the model posits that whereas alcohol's effects on emotional response are indirect, alcohol does yield several consistent effects on certain aspects of cognitive processing: Alcohol both reduces overall attentional processing capacity (working memory) and induces a state of attentional narrowing, such that after drinking, individuals cannot attend to as many stimuli as they could prior to drinking. The ramifications of the theory suggest that if, for example, an individual who is anxious about an upcoming examination drinks in the presence of benign distraction (e.g. seated at a local bar with friends and music playing), his narrowed and shrunken attention would necessarily be reallocated to the more salient distraction and away from the cognitions that were otherwise promoting the anxiety. Under these circumstances, drinking in combination with such pleasant distracters would result in anxiety reduction. If, however, that same individual chooses to drink at home, alone in his kitchen (ostensibly in the absence of pleasant distraction), his attention would remain focused on the anxious cognitions and, hence, not only would consumption of alcohol not serve to reduce anxiety, it would exacerbate it.

Kassel and colleagues applied the basic tenets of this model to better understanding cigarette smoking's seemingly inconsistent effects on stress and anxiety reduction. Their findings were remarkably similar to those observed with alcohol: Whereas anxious smokers

who smoked in the presence of benign distraction experienced dramatic decreases in self-reported anxiety, those smokers who smoked in the absence of distraction experienced either no change or an increase in anxiety. Thus, the model appears to offer an interesting framework for understanding, and even predicting, emotional response as influenced by two very different drugs: alcohol and nicotine. Nonetheless, more research is needed to determine whether this proposed mechanism – reallocation of attentional resources – is volitional or governed by forces operating outside of conscious awareness.

Appraisal Disruption Model

The appraisal disruption model rests on the assumption that before experiencing an emotional response, cognitive processing must occur within the individual in order to evaluate and appraise the subjective meaning (i.e. personal relevance) of the antecedent causal event (e.g. argument with spouse). Thus, it is the appraisal of the situation that produces one's emotional response. If this cognitive appraisal is disrupted, an emotional response would be lessened or blunted as the information normally primed and triggered to formulate emotion would not be accessible in memory. Many substances, including alcohol and tobacco, have been found to disrupt the cognitive faculties critical to this appraisal process (e.g. alcohol reduces response to an anxiety-provoking situation; going on a first date). Substance use, particularly alcohol, is thus believed to influence emotion through this particular cognitive pathway.

ADDICTION AND EMOTION: DEVELOPMENTAL PHASES

Thus far, we have discussed the burgeoning evidence pointing to robust associations between drug addiction and various indices of emotion. Moreover, these links are steeped in sound theoretical frameworks. We now consider the notion that the relationship between substance abuse, addiction, and emotion may vary across developmental stages of addiction. Hence, we briefly review such associations across the phases of drug initiation, maintenance, and relapse.

Initiation

In recent years, research has increasingly focused on the role played by affective states, and their interaction with the environment, in the initiation of substance use and the development of subsequent disorders (i.e. addiction). When considering emotion regulation, a compelling area of interest is the transition from adolescence to adulthood, sometimes referred to as emerging or young adulthood. These developmental stages are when most addicts begin using drugs. This period is marked by a general increase in NA and emotional lability, as well as a widening array of social stressors. To deal with these stressors, adolescents and young adults typically use emotion-focused and/or problem-focused coping strategies. Emotion-focused coping strategies manage the emotions linked to, or resulting from, stressors over which individuals believe they have little or no control. Problem-focused strategies, however, focus on effectively altering or removing the stressor and attendant feelings.

Affective dysregulation, or poor emotional control, and routine use of emotion-focused strategies have been linked to increased alcohol and substance use. When the drug itself is used to manage emotion, the use itself becomes an emotion-focused coping strategy. Thus, just as NA increases during young adulthood, so do rates of alcohol use and abuse. This is especially true in the years immediately following high school, before these rates peak when adults reach their early 20s.

During this time, alcohol expectancies play a large role in motivating drinking behavior. Drinking for positive reinforcement (i.e. to get high) is associated with positive alcohol expectancies and overall reward seeking. As discussed earlier in this chapter, these motivations are most likely based on alcohol's activation of reward centers in the brain. Conversely, others report drinking to relieve NA (i.e. to forget about problems) and have a very strong belief that alcohol reduces stress and anxiety. Longitudinal research strongly suggests that those adolescents and young adults whose drinking is primarily motivated by negative reinforcement are at much greater risk of subsequently developing abuse and addiction problems. Regarding the extent to which alcohol (and other drugs of abuse) actually serve to reliably reduce NA and other indices of affective distress among adolescents and young adults, the jury is still out. It appears that the effects of alcohol and other drugs on NA are highly conditional, relying on both individual (e.g. expectancies) and environmental factors.

Alcohol use during adolescence and emerging adulthood predicts both concurrent and later tobacco and illicit substance use. Accordingly, most experimental smoking occurs during adolescence; the prevalence of all substance use disorders increases steadily throughout this time period. The extant literature indicates that, like many adult smokers, a significant proportion of adolescents smoke initially to relieve stress and NA. Critically, longitudinal studies of adolescents have shown that those who become regular smokers perceive more stress in their lives and do not use coping strategies effectively, relative to those whose smoking does

not progress. Hence, for these at-risk adolescents, smoking is more typically viewed as a viable coping resource.

Maintenance

Emotional distress plays a crucial role in the transition from controlled to excessive substance use. Persistent substance use often represents a way to cope with negative emotions and obtain rewards, as acute substance use can result in both positive arousal and alleviation of NA. It follows that addicted persons often experience heightened negative emotional states, prompting persistent, compulsive substance use without thought of negative consequences.

The self-medication hypothesis, as originally informed by psychoanalytic theory, but later modified by social–cognitive perspectives, suggests that continued substance use functions as a means to modulate affect and soothe psychological distress. Accordingly, it asserts that substance users find negative emotions intolerable and overwhelming and are unable to manage these emotional states on their own. Thus, they seek the physiological and psychological effects of their chosen substance to regulate their distress and achieve emotional stability. Whereas this model has garnered empirical support, the specific emotional effects of individual substances remain unclear. Hence, as discussed earlier in the section on Theoretical Frameworks, it appears that contextual factors likely play an important role in modifying drug effects on emotional response.

An alternate tripartite model helps explain how substance use is maintained, describing the central components of the addiction cycle in clinical terms: preoccupation–anticipation, binge–intoxication, and withdrawal–NA. First, there is an initial preoccupation with obtaining the substance despite persistent physical or psychological consequences. Subsequent intoxication results in the failure to fulfill social, occupational, and recreational obligations. These failures often result in increased NA that is exacerbated by emerging withdrawal symptoms. A cycle of repeated failures to self-regulate follows, and each violation brings greater NA and results in greater tolerance than the last. A similar model of addiction asserts that the relief of NA is the primary motivation for substance use. Through repeated cycles of substance use and withdrawal, users become more sensitive to the interoceptive cues of NA and the soothing effects of the substance.

It follows that negative emotionality, anxiety, and depression have all been linked to alcohol use and other substance use disorders. For example, compared with social drinkers, excessive users often report drinking to relieve personal distress and cope with negative emotions. Those with alcohol use disorders manifest higher rates of anxiety and depression and score higher on measures of neuroticism and negative emotionality than nondisordered individuals. The relationship between emotional distress, emotion regulation, and drinking is complex, however, as prior research identifies NA as both a predictor and consequence of alcohol use, creating a cycle of negative emotionality. Self-regulation, the process by which one formulates a plan and adjusts one's behavior accordingly, is an important component in managing this kind of emotional response. Deficits in self-regulation have been reliably linked to the development and maintenance of substance use disorders. The most effective interventions often focus on increasing self-regulation to break this affective cycle.

Correspondingly, although there is ample anecdotal evidence of smoking's ability to relieve NA among regular smokers, research findings have been mixed. Generally, the psychological gains experienced while smoking are thought to result from the reversal of withdrawal symptoms. Indeed, withdrawal effects have been demonstrated with increasing accuracy. Researchers have begun to measure mood before and after participants smoke a series of cigarettes. Mood ratings generally improve after they smoke the first cigarette compared with baseline, decline during the abstinence period, and improve again after they smoke the second cigarette. Similar results have been shown with self-reported stress and arousal.

Relapse

As discussed throughout this chapter, most models of drug use identify negative mood states, such as stress and depression, as important precipitants and motivators of drug use. Moreover, drug abuse and dependence have been conceptualized as chronically relapsing conditions. Decreased NA as well as an increase in positive affect may both play a role in continued drug use, including vulnerability to relapse. The relationship between NA and relapse, perhaps even more so than NA's relationship with use during the initiation and maintenance phases of use, appears particularly robust.

One way to view the relationship between emotion and drug relapse is through a conditioning paradigm. Over time, chronic drug users may experience negative mood states, such as stress and depression, as conditioned stimuli that illicit craving for the drug. Several laboratory studies have shown increased craving among drug users after exposure to NA cues, such as visual images or music. Animal studies have also indicated that stress often leads to reinstated drug-seeking behavior among abstinent animals. Moreover, reinstatement of drug-seeking behavior in animals, in response to drug cues as well as

stress cues, has been shown to increase dopamine release in the nucleus accumbens. Therefore, the conditioning properties of both positive and negative cues have been attributed to dopamine release in this brain region.

Another model in which emotion and relapse have been conceptualized is through proposed brain changes in reward and stress systems. Most drugs of abuse stimulate the mesocorticolimbic dopaminergic systems, the brain reward pathways, accounting for the rewarding properties of drugs of abuse. Furthermore, neural stress and reward pathways are closely related. Therefore, stress pathways in the brain, including norepinephrine systems and corticotrophin-releasing factor, are also affected by drug use. Several theorists postulate that neural changes to these brain reward and stress systems caused by repeated drug use may increase vulnerability to stress and in turn, increase the risk of relapse.

Robust evidence exists for the relationship between stress and relapse to several drugs including nicotine, alcohol, heroin, and cocaine. However, not all abstinent drug users who experience stressful life events relapse. Furthermore, it appears that some individuals are better able to cope with stressful life events compared with others. Several studies have shown that the use of adaptive coping mechanisms is associated with decreased risk of relapse. Both the type of coping mechanism and the number of coping mechanisms used have been shown to be important factors in predicting relapse. However, as noted previously, chronic drug use may lead to neuroadaptations in the stress system. Thus, not only may drug users possess an increased sensitivity to stress, they may also be less well equipped to cope with stress. Some have proposed that teaching coping mechanisms to recovering drug users should represent a critical component of any drug and alcohol intervention.

CONCLUSIONS AND FUTURE DIRECTIONS

Based on a review of the empirical literature, we have argued that the relationship between drug addiction and emotional processes is a critical one. The idea that people self-administer drugs and alcohol because of their respective affective payoffs is hardly new. It has endured for centuries, is steeped in common folk psychology, and certainly holds intuitive appeal. None of these realities reify the notion that drug use and abuse is integrally tied to emotional processes. In order to truly understand and assess the links between emotion and drug addiction, science must turn to rigorous studies rooted in sound theoretical bases. Fortunately, in recent years, we believe that the field has been moving in such directions.

At the level of large-scale epidemiological inquiry, the relationship between drug use and emotional distress, as expressed both in subclinical symptoms and full-blown disorders, is borne out again and again. The question of directionality remains at least partially unanswered, although burgeoning evidence seems to support bidirectionality: drug use predisposes to affective distress, and affective distress renders individuals vulnerable to drug abuse and addiction. A thorough understanding of such relationships does have bearing in very real ways on both prevention and intervention approaches to substance abuse and addiction. It would be difficult to find any state-of-the-art addiction programs that do not incorporate some aspect of emotion regulation, affect tolerance, and/or adaptive coping into their treatment regimens.

We conclude this chapter with five recommendations, or considerations, for those interested in further examining drug/emotion associations.

(1) Whereas cross-sectional epidemiological surveys are quite useful in establishing associations, at the population level, among various indices of affective distress and drug-related problems, one must be careful not to infer causal processes from such data. So, for example, whereas it is compelling to observe that depression in adolescence significantly heightens vulnerability to drug abuse and addiction in young adulthood and adulthood, such data do not indicate causal mechanisms (e.g. that depression causes addiction or that depressed people self-administer drugs in order to cope with their depression). Whereas such inferences may turn out to be valid, they must be scientifically addressed in their own right.

(2) The study of drug/emotion associations must be steeped in sound theoretical frameworks. In this chapter, we provided brief overviews of several influential theories that have bearing on this issue. This overview is not complete, however, and many other sound theoretical perspectives are available to guide research. Our concern is that research in this area frequently uses non-theoretical approaches, and that this state of affairs ultimately renders the field's database and understanding of these important issues weaker by decree.

(3) In this chapter, we alluded to the likelihood that drug/emotion relationships change over the developmental course of drug use. Hence, the association between, say, NA and cigarette smoking may be different in the initiation phase than it is in the maintenance phase (when smokers have escalated to daily dependent use). An important implication of this suggestion is that the stage of drug use must be taken into account when exploring the association between use and emotion.

(4) The links between drug addiction and emotion are likely moderated by a host of variables. Thus, future

research in this area should consider and assess the potential role played by such moderating factors. Likely candidate variables include gender, cultural factors, personality variables (e.g. neuroticism, sensation seeking, impulsivity), and a host of contextual variables.

(5) We believe it is important to note that drug researchers frequently reside in independent camps, resulting in a lack of communication between those who study, for example, cigarette smoking, or heroin use, or cocaine use, or alcohol use, and so on. Although it may turn out that the associations between the use of these respective drugs to emotional processes are unique to that particular drug, it would seem that the field would benefit tremendously from more cross talk among researchers ensconced in their particular drug camps.

In summary, the field of addiction science has made tremendous progress in recent years, both in describing and understanding the causative processes governing the development of drug addiction, as well as applying such knowledge to treatment programs. But much work lies ahead. We believe that further study of the crucial links between drug addiction and emotional processes will yield tremendous benefits in terms of lives saved.

Acknowledgments

Preparation of this chapter was made possible, in part, by the National Cancer Institute (grant 5PO1 CA098262).

SEE ALSO

The Biopsychosocial Model of Addiction, Contextual Factors in Addiction, Epidemiology of Addiction, Personality and Addiction Processes, Relapse and Lapse, Self-Medication, Substance Induced Myopia, The Terminology of Addictive Behavior, Deprivation, Craving, and Affect: Intersecting Constructs in Addiction, Craving and Expectancies, Substance Use and Mood Disorders

List of Abbreviations

NA	negative affect
RMNR	Reformulated Model of Negative Reinforcement
UCR	unconditioned response
UCS	unconditioned stimulus

Further Reading

Al'Absi, M. (Ed.), 2007. Stress and Addiction: Biological and Psychological Mechanisms. Academic Press, Burlington, MA.

Audrain-McGovern, J., Rodriguez, D., Kassel, J.D., 2009. Adolescent smoking and depression: evidence for self-medication and peer smoking mediation. Addiction 104, 1743–1756.

Baker, T.B., Piper, M.E., McCarthy, D.E., Majeskie, M.R., Fiore, M.C., 2004. Addiction motivation reformulated: an affective processing model of negative reinforcement. Psychological Review 111, 33–51.

Berridge, K.C., Robinson, T.E., Aldridge, J.W., 2009. Dissecting components of reward: "liking," "wanting," and learning. Current Opinion in Pharmacology 9, 64–73.

Earleywine, M. (Ed.), 2005. Mind-altering Drugs: The Science of Subjective Experience. Oxford University Press, New York.

Kassel, J.D. (Ed.), 2010. Substance Abuse and Emotion. American Psychological Association, Washington, DC.

Kassel, J.D., Evatt, D.P., Greenstein, J.E., Wardle, M.C., Yates, M.C., Veilleux, J.C., 2007. The acute effects of nicotine on positive and negative affect in adolescent smokers. Journal of Abnormal Psychology 116, 543–553.

Kassel, J.D., Heinz, A.J., Evatt, D.P., Braun, A.R., 2010. Treatment models for comorbid psychiatric and addictive disorders. In: Miller, N.S., Gold, M.S. (Eds.), Addictive Disorders in Medical Populations. John Wiley and Sons, Hoboken, NJ.

Kassel, J.D., Stroud, L., Patronis, C., 2003. Smoking, nicotine, and stress: correlation, causation, and context across stages of smoking. Psychological Bulletin 129, 270–304.

Koob, G.F., Le Moal, M., 2008. Addiction and the brain antireward system. Annual Review of Psychology 59, 29–53.

Leonard, K.E., Blane, H.E. (Eds.), 1999. Psychological Theories of Drinking and Alcoholism. Guilford Press, New York.

Panksepp, J., 1998. Affective Neuroscience: The Foundation of Human and Animal Emotions. Oxford University Press, New York.

CHAPTER

23

Stress and Addiction

Rajita Sinha

Yale University School of Medicine, New Haven, CT, USA

OUTLINE

INTRODUCTION

Stress is known to increase the risk of addictive behaviors. The last two decades have led to a dramatic rise in research on understanding the underlying mechanisms for this association. Behavioral and neurobiological correlates are being identified and there is growing evidence of molecular and cellular changes associated with chronic stress and addiction. Human studies have benefited from the emergence of sophisticated brain imaging tools and the cross-examination of laboratory-induced methods of stress and craving and their association with specific brain regions associated with reward and addiction risk. This chapter focuses on these links between stress and addiction in humans, but also draws on supporting evidence from the broader animal literature. A definition of stress and its neural underpinnings are presented with specific emphasis on its effects on motivation and behavior. Epidemiological evidence linking early childhood and adult adversity with risk of addiction, and current research on the putative mechanisms underlying this association is presented. The prefrontal brain circuits are seen as playing a critical role in adaptive learning and executive function, including controlling distress and desires/impulses, and in the association between stress and addiction risk. The effects of chronic drug use on brain reward and motivation pathways combined with the premorbid

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00023-1

and ongoing effects of repeated and chronic stress are presented as the pathophysiology that promotes addiction relapse risk. Finally, novel prevention and treatment approaches to addressing the deleterious effects of stress, and the combined effects of stress and addictive behaviors on relapse are discussed.

STRESS, EMOTIONS, AND ADAPTIVE BEHAVIOR

Definition and Conceptualization of Stress

The term "stress" refers to processes involving perception, appraisal, and response to harmful, threatening, challenging events or stimuli. Stressors are stimuli that are overwhelming and challenging to an individual. Stress experiences can be emotionally or physiologically challenging, and they activate stress responses and adaptive processes to regain homeostasis. Examples of emotional stressors include interpersonal conflict, loss of relationship, death of a close family member, and loss of a child. Highly threatening or stressful events also include emotional and physical, traumas such as violent victimization and emotional, physical, and sexual abuse. Common physiological stressors are hunger or food deprivation, sleep deprivation or insomnia, psychoactive drug use, extreme hyper or hypothermia, and drug withdrawal states. This kind of conceptualization allows for consideration of (1) internal and external events or stimuli that exert demands or load on the individual; (2) the neural processes that evaluate the demands and assess availability of adaptive resources to cope with the demands (appraisal); (3) the subjective, behavioral, and physiological activity that signal stress to the organism; (4) neural and physiological adaptations in emotional and motivational systems associated with repeated and/or chronic stress; and (5) behavioral, cognitive, and physiological adaptation in response to stressors.

While stress is often associated with negative emotions, external and internal stimuli that are mild/moderately challenging but limited in duration, and result in flexible cognitive, physiological, and behavioral responses with a sense of mastery and accomplishment can be perceived as pleasant and/or, exciting. Such experiences are often thought of as opportunities for learning and growth, particularly in the context of resilience factors that promote mastery (e.g. social support, adequate motivational and executive functioning), to achieve goal-directed outcomes and homeostasis. On the other hand, the more prolonged, repeated or chronic the stress (e.g. states associated with increased intensity or persistence of distress), the greater the uncontrollability and unpredictability of the stressful situation, lower the sense of mastery, greater the risk of inflexible behavioral and physiological responses, and greater the magnitude and persistence of the stress response, with higher risk of the negative effects on brain and body systems involved in adaptive function. Thus, the dimensions of intensity, controllability, predictability, mastery, or adaptability are important in understanding the role of stress in increasing risk of maladaptive behaviors such as addiction.

Neural and Physiological Pathways Involved in the Stress Response

The perception and appraisal of stress rely on specific aspects of the presenting external or internal stimuli, personality traits, availability of internal resources, physiological condition of the individual, emotional state, including beliefs and expectancies and specific neural state of brain regions mediating the appraisal of stimuli as distressing and the resulting physiological, behavioral, and emotional experiences and adaptive responses. The brain cortical regions such as the insular and prefrontal cortices along with the middle temporal regions play a key role in perceiving and appraising emotion and stress states. Also, the brain's limbic and striatal regions contribute to identifying emotional signals, learning about emotional stimuli and their aversive and rewarding properties and contribute to the emotional experience of stress and emotional stimuli. The brainstem (locus ceruleus and related arousal regions), hypothalamus, thalamus, striatal and limbic regions are also involved in the physiological, emotional, and behavioral responses and together these regions contribute to the experience of distress.

The physiological responses are manifested through two major stress systems, namely, the hypothalamic-pituitary-adrenal (HPA) axis and the autonomic nervous system, which in turn also influence the immune system. These stress pathways are intricately connected with each other and also with the brain pathways described above so that a coordinated stress response may be mobilized in the face of challenge. The HPA axis is stimulated by the corticotropin releasing factor (CRF) released from the paraventricular nucleus (PVN) of the hypothalamus, to stimulate the adrenocorticotropin hormone from the anterior pituitary, which initiates the secretion of cortisol/corticosterone from the adrenal glands. The autonomic nervous system includes the sympathetic nervous system that mobilizes the cardiovascular and immune arousal responses and the parasympathetic nervous system that is involved in regulating the sympathetic arousal by providing the "brakes" to the sympathetic arousal and in regaining homeostasis via the sympathoadrenal medulary systems. Recent evidence

also indicates that the sympathetic pathways provide further modulation of the adrenal glands for release of cortisol as well as release of norepinephrine (NE) and epinephrine (EPI). These are the core stress pathways involved in stress arousal and mobilization of the body and brain to respond to stress and in regulation of stress so as to regain homeostasis.

In addition, CRF has extensive influence in extrahypothalamic regions across the corticostriatal-limbic pathways and plays a critical role in modulating the subjective and behavioral stress responses. Furthermore, central catecholamines, particularly noradrenaline and dopamine, are involved in modulating brain motivational pathways (including the ventral tegmental area (VTA), nucleus accumbens (NAc), and the medial prefrontal cortex (mPFC) regions) that are important in regulating distress, exerting cognitive and behavioral control, and influencing choice or response selection and decision making critical for adaptation and homeostasis. The hypothalamic and extrahypothalamic CRF pathways and central catecholamines target brain motivational pathways to critically affect adaptive and homeostatic processes. For example, different parts of the mPFC are involved in higher cognitive or executive control functions, such as controlling and inhibiting impulses, regulating distress, focusing and shifting attention, working memory, monitoring conflict and behavior, linking behaviors and consequences over time, and considering alternatives before acting and decision making responses. Psychosocial and behavioral scientists have elegantly shown that with increasing levels of emotional and physiological stress or negative affect, there is a decrease in working memory function, poor attention and flexibility, lower behavioral control and increases in impulsive responding.

Neurobiological Responses to High Uncontrollable Stress

Recent evidence from human brain imaging research shows that recent life stressors and chronic stress decreases gray matter volume in medial prefrontal and hippocampus and insula regions of the brain. These are key regions in perception and appraisal of stress and emotions and in regulating the physiological and behavioral response to stress. Similarly, exposure to recent life stress and acute stress decreases responses in the dorsolateral and mPFC associated with working memory and reward processing. Thus, with increasing levels of stress, there is a decrease in prefrontal functioning and increased limbic-striatal level responding, a brain pattern associated with low behavioral and cognitive control. Low behavioral and cognitive control linked to the prefrontal and insular cortex and high responding in limbic-emotional and striatal-motivation brain regions under stress provides the specific pattern for promoting addictive behavioral patterns where there is a decreased ability to control rewarding behaviors. Thus, motivational brain pathways are key targets of stress chemicals which point to an important potential mechanism by which stress affects addiction vulnerability.

STRESS AND THE DEVELOPMENT OF ADDICTIVE BEHAVIORS

A substantial amount of literature exists on the significant association between acute and chronic stress and the motivation to abuse addictive substances. Many of the major theories of addiction also identify an important role of stress in addiction. These range from psychological models that view drug use and abuse as a coping strategy to deal with stress, to reduce tension, to self medicate, and to decrease withdrawal-related distress, to neurobiological models that propose incentive sensitization and stress allostasis concepts to explain how neuroadaptations in reward, learning and stress pathways may enhance the key features of addiction, namely, craving, loss of control, and compulsion. This section reviews the converging lines of evidence that point to the critical role that stress plays in increasing addiction vulnerability.

Chronic Stress and Adversity and Increased Vulnerability to Drug Use

Considerable evidence from population-based and clinical studies indicate a positive association between psychosocial adversity, negative effect, and chronic distress and addiction vulnerability. It is important to note that there are some negative studies that do not support the notion that stress increases drug use, and such studies highlight the need to consider a variety of measurement issues, such as, measurement of stress (e.g. acute versus chronic; subjective or perceived stress versus repeated adverse life events, mood versus stress induction, subjective rating of stress) and measurement of addictive behaviors (self-report, objective tests; assessment of one substance such as alcohol or multiple substances including nicotine) and measurement over a period of days (single assessment versus repeated real life daily assessment, retrospective reporting versus prospective reports). Despite these caveats, the majority of studies show positive support for the association between stress and addictive behaviors. The evidence in this area can be categorized into four broad types. The first are prospective studies demonstrating that

adolescents facing high recent negative life events show increased levels of drug use and abuse. Negative life events such as loss of parent, parental divorce and conflict, low parental support, physical violence and abuse, emotional abuse and neglect, isolation and deviant affiliation, and single parent family structure have all been associated with increased risk of subsequent substance abuse.

The second are data showing the association between trauma and maltreatment and addiction, and negative affect, chronic distress, and risk of substance abuse. Overwhelming evidence exists for an increased association between childhood sexual and physical abuse, victimization and increased drug use and abuse. In addition to sexual and physical abuse, negative effect and chronic distress states are predictive of addiction vulnerability. Findings indicate that negative effect, including temperamental negative emotionality are associated with substance abuse risk. Several studies have also shown a significant association between prevalence of mood and anxiety disorders, including post-traumatic stress disorder, behavioral conduct problems, and increased risk of substance use disorders.

The third set comes from research on lifetime exposure to stressors and the impact of cumulative adversity on addiction vulnerability after accounting for a number of control factors such as race/ethnicity, gender, socioeconomic status, prior drug abuse, prevalence of psychiatric disorders, family history of substance use, and behavioral and conduct problems. Increasing levels of cumulative stressful life events are significantly predictive of alcohol and drug dependence in a dose-dependent manner, even after accounting for control factors. Furthermore, the dose-dependent effects of cumulative stressors on risk for addiction exists for both genders and for Caucasian, African-American, and Hispanic race/ethnic groups. The types of adverse events significantly associated with addiction vulnerability are parental divorce or conflict, abandonment, forced to live apart from parents, loss of child by death or removal, unfaithfulness of significant other, loss of home to natural disaster, death of a close one, emotional abuse or neglect, sexual abuse, rape, physical abuse by parent, caretaker, family member, spouse or significant other, victim of gun shooting or other violent acts, and observing violent victimization. These represent highly stressful and emotionally distressing events, which are typically uncontrollable and unpredictable in nature (see Table 23.1).

Finally, the fourth line of evidence comes from human laboratory experiments and studies using ecological momentary assessment techniques to assess ongoing drug use in daily life. Laboratory studies provoke emotional, cognitive, or social stress in the context of an experiment followed by opportunities for drug self-administration. Such studies examining effects of stress exposure on drug use are limited to legal drugs such as alcohol and nicotine, due to ethical reasons of providing opportunities for illicit drug use to nonaddicted or nonexposed samples. Findings from these studies generally support the notion that acute exposure to stress or negative mood or high ratings of negative mood and stress are associated with alcohol and nicotine use, but the effects of drinking history, history of adversity, social stress and expectancies are known to play a role in these studies.

Putative Mechanisms Underlying Stress Effects on Addiction Vulnerability

This section presents the neurobiological links between stress and reward pathways activated by abusive drugs. It is well-known that the reinforcing properties of drugs of abuse involve activation of the mesolimbic dopaminergic (DA) pathways, which include dopamine neurons originating in the VTA and extending to the ventral striatum (VS/NAc) and the PFC. This pathway is also involved in assigning salience to stimuli, in reward processing, and in learning and adaptation. Human brain imaging studies support the role of these systems in drug reward, as psychostimulants, alcohol, opioids, and nicotine all activate the mesolimbic DA systems, in particular, the ventral and dorsal striatum, and such activity has been associated with the drug ratings of high/euphoria and/or craving.

Interestingly, biological stress responses with increases in the stress hormone, cortisol (corticosterone in animals) can significantly affect dopamine signaling in the reward and motivation pathways. For example, stress exposure and increased corticosterone enhance dopamine release in the NAc. Suppression of corticosterone by surgical removal of the adrenal glands (adrenalectomy) reduces extracellular levels of dopamine under basal conditions and in response to stress and psychostimulants. However, chronic corticosterone inhibits DA synthesis and turnover in the NAc. There is also evidence that like drugs of abuse, stress and concomitant increases in CRF and corticosterone enhance the activity of the excitatory neurotransmitter glutamate, in the VTA, which in turn, enhances activity of DA neurons. Human brain imaging studies have further shown that stress-related increases in cortisol are associated with increased dopamine transmission in the VS, and some evidence also reveals that amphetamine-induced increases in cortisol are associated with both dopamine binding in the VS and with ratings of amphetamine-induced euphoria. Furthermore, both stress and drugs of abuse activate the mesolimbic pathways, and

TABLE 23.1 Types of Adverse Life Events, Trauma, Chronic Stressors, and Individual-level Variables Predictive of Addiction Risk

Adverse Life Events and Chronic Stressors	Life Trauma	Individual-level Variables
- Loss of parent - Parental divorce and conflict - Isolation and abandonment - Single parent family structure - Forced to live apart from parents - Loss of child by death or removal - Unfaithfulness of significant other - Loss of home to natural disaster - Death of significant other/close family member - Victim of gun shooting or other violent acts - Observing violent victimization	- Physical neglect - Physical abuse by parent/caretaker/family member/spouse/significant other - Emotional abuse and neglect - Sexual abuse - Rape	- Negative emotionality - Poor behavioral control - Poor emotional control

Reprinted with permission from Sinha (2008).

each result in synaptic adaptations in VTA dopamine neurons and in morphological changes in the mPFC.

In addition to a role in reward, a growing body of human imaging studies and preclinical data indicate that the VS NAc is also involved in aversive conditioning, experience of aversive, pain stimuli, and in anticipation of aversive stimuli. Such evidence points to a role for the mesolimbic dopamine pathways beyond reward processing, and one that more broadly involves habit formation, motivation, and attention to behavioral response during salient (aversive or appetitive) events. Furthermore, additional regions connected to the mesolimbic DA pathways and involved in reward, learning and adaptive, goal-directed behaviors are the amygdala, hippocampus, insula, and related corticolimbic regions. These regions along with the mesolimbic DA pathways play an important role in interoception, emotions and stress processing, impulse control and decision making, and in the addictive properties of drugs of abuse.

Stress Mechanisms Involved in Acquisition of Drug Self-administration

Animal studies have shown that stress exposure and corticosterone administration increase the rapidity with which animals learn to self-administer abusive drugs. Corticosterone administration facilitates psychomotor stimulant effects of cocaine and morphine. Furthermore, antagonists of the glucocorticoid (GC) receptors to which corticosterone (and cortisol) binds, injected into the VTA decreases morphine-induced locomotor activity, suggesting that activity of GC receptors in the VTA could mediate dopamine-dependent behavioral effects. Mice with deletion of the GC receptor gene shows a dose-dependent decrease in motivation to self-administer cocaine. These data suggest that HPA-related corticosterone release could partially mediate the dopamine increase soon after drug administration.

Although in nonhuman primates the link between cortisol, dopamine, and drug self-administration has not been reported, there is evidence that stress related to social subordination is associated with lower levels of D2 receptors and higher cocaine self-administration. In humans, Positive Emission Tomography (PET) studies using [^{11}C]raclopride indicate that acute stress exposure increases dopamine release in the VS. For example, healthy individuals with low early life maternal care showed greater dopamine release in the VS during an acute psychological stress task as compared to those with a history of high early life maternal care. Furthermore, cortisol response during the stress task correlated significantly with VS dopamine release. Although these data support the link between stress/cortisol and dopamine transmission, human research linking stress-induced changes in VS activity or dopamine binding and risk of addictive behavior is needed to directly establish the association between stress, mesolimbic dopamine, and addiction risk.

Early Life Chronic Stress, Dopamine Systems, and Drug Self-administration

There is growing evidence from basic science studies that early life stress and chronic stress significantly affect the mesolimbic dopamine pathways and play a role in drug self-administration. Repeated and prolonged exposure to maternal separation in neonatal rats significantly alters the development of central CRF pathways. These animals as adults show exaggerated HPA and

behavioral responses to stress. Such physiological and behavioral changes are associated with altered CRF mRNA expression in the PVN, increased CRF-like immunoreactivity in the locus coeruleus (LC), and increased CRF receptor levels in the LC and raphe nuclei. The adult animals also show decreased negative feedback sensitivity to GC and these changes are accompanied by decreased GC receptor expression in the hippocampus and frontal cortex. Decreased gamma-aminobutyric acid (GABA) receptor levels in noradrenergic cell body regions in the LC and decreased central benzodiazepine receptor levels in the LC and the amygdala have also been reported. More importantly, rats exposed to maternal separation in childhood show significantly elevated DA responses to acute stress along with increased stress-induced behavioral sensitization and robust behavioral sensitization to psychostimulant administration, suggesting heightened behavioral responses to psychostimulants. Such cross-sensitization of stress and drugs of abuse is associated with enhanced release of DA in the NAc, lower NAc-core, and striatal DA transporter sites, and reduced D3 receptor binding sites and mRNA levels in the NAc shell. In addition, changes in brain NE are also known to alter DA signaling and increase behavioral and psychostimulant sensitization.

A number of studies have shown that early life stress and prolonged and repeated stress adversely affects development of the PFC, a region that is highly dependent on environmental experiences for maturation. Repeated and chronic stress-related changes in prefrontal cortical neuron spines, density, and synapses have been documented in animal studies. The PFC, and particularly the right PFC, plays an important role both in activating the HPA axis and autonomic responses to stress and in regulating these responses. For example, lesions of the ventromedial PFC result in enhanced HPA and autonomic responses to stress. High levels of GC receptors are also found in the PFC and chronic GC treatment results in a dramatic dendritic reorganization of PFC and hippocampal neurons. Human studies on the neurobiological effects of child maltreatment document neuroendocrine changes as well as alterations in size and volume of prefrontal, thalamic, and cerebellar regions associated with maltreatment and with initiation of addiction. Similarly, adult human structural neuroimaging studies show stress-related decrease in gray matter volume in the medial PFC and reduced gray matter volume in prefrontal and anterior cingulate regions in samples of addicted populations. Together, the data presented in this section highlight the significance of stress effects on mesolimbic and prefrontal regions that are important in stress regulation, behavioral control, and decision making.

Stress, Self-control, and Addiction Vulnerability

High uncontrollable stress is associated with loss of control over impulses, an inability to inhibit inappropriate behaviors and to delay gratification. Neurobiological data indicate that stress impairs catecholamine modulation of prefrontal circuits, which in turn impairs executive functions like working memory and self-control. There is also growing evidence that adolescents with high exposure to life stress (as those listed in Table 23.1) and at risk for substance abuse are more likely to show decreased emotional and behavioral control, and decreased self-control is associated with risk of substance abuse and other maladaptive behaviors. Adolescents at risk for substance abuse are known to have decreased executive functioning, low behavioral and emotional control, poor decision making, and greater levels of deviant behavior and impulsivity. The corticostriatal-limbic dopamine pathways have been associated with impulsivity and decision making and addiction risk, and as discussed in previous sections, specific regions of this pathway, such as the VTA, NAc, PFC, and amygdala, are highly susceptible to stress-related signaling and plasticity associated with early life stress, repeated and chronic stress experiences. Interactions between chronic stress and impulsivity and dopamine responses in the VS have also been observed. Individuals with low to moderate stress and low impulsivity had greater dopamine release, but in individuals with high stress, both low and high impulsive individuals showed low DA release in response to amphetamine. These findings demonstrate the important effects of stress and impulsivity on mesolimbic dopamine transmission, and highlight that both factors need to be carefully considered to fully understand the role of stress and impulsivity on addiction risk.

The above sections are summarized in a schematic diagram in Fig. 23.1 to illustrate how stress may affect addiction vulnerability. It highlights cross-sensitization of stress and drug abuse on specific behavioral and neurochemical responses, and indicates the common neurobiological pathways upon which both stress and drugs of abuse act (Fig. 23.1 reprinted from Sinha, 2008).

DRUG USE AND ABUSE AND CHANGES IN STRESS AND REWARD PATHWAYS

There is growing evidence that abusive drugs when taken acutely and chronically potently influence brain stress pathways. These changes in turn alter cognitive, subjective, and behavioral responses to stress and also

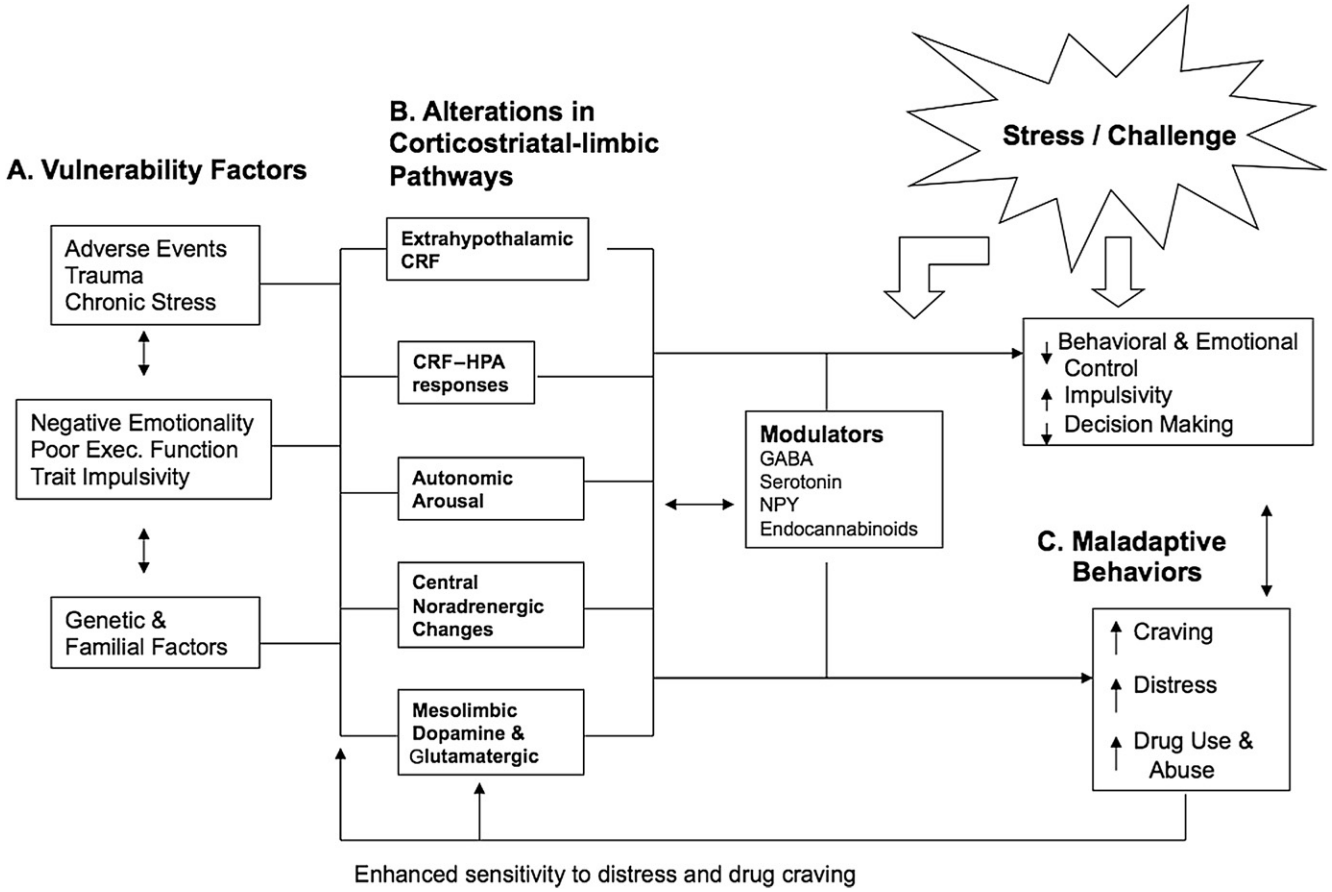


FIGURE 23.1 A schematic model of stress effects on addiction representing the cross-sensitization of stress and drugs on behavioral and neurochemical responses that are mediated by the stress and reward pathways. Column A lists three types of vulnerability factors: (1) developmental/individual-level factors such as frontal executive function development, negative emotionality, behavioral/self-control, impulsivity or risk taking, altered initial sensitivity to rewarding effects of drugs; (2) stress-related vulnerability factors such as early adverse life events, trauma and child maltreatment experiences, prolonged and chronic stress experiences; and (3) genetic influences and family history of psychopathology and addiction which have not been discussed here but have significant interactive effects on addiction risk and in emotion and stress markers. Each of these factors may influence each other to significantly affect alterations in neurobiological pathways involved in stress regulation and cognitive and behavioral control (column B). Specific synaptic changes in these pathways at molecular and cellular levels provide the basis for the mechanism by which stress and individual and genetic factors in column A interact to increase risk of maladaptive behaviors represented in column C. The model suggests that stress experiences in the presence of these vulnerability factors result in maladaptive stress and self-control responses that increase addiction risk. The specific mechanism by which the maladaptive stress responding increases this risk involve dysregulation in brain stress circuits, particularly the CRF and NE systems, and their interactions with the mesocorticolimbic-striatal dopamine pathways and its modulation by glutamate and GABA. Furthermore, recent evidence suggests that stress regulatory molecules, including neuropeptides such as neuropeptide Y, endocannabinoids, brain-derived neurotrophic factor (BDNF), and neuroactive steroids play a role in addiction vulnerability. *Reprinted with permission from Sinha (2008).*

impact desire/craving for drug and perpetuations of addictive behaviors.

Acute and Chronic Drug Use and Changes in Stress Responses

It is well-known that acute administration of the most commonly abused drugs, such as alcohol, nicotine, cocaine, amphetamines, and marijuana that activate brain reward pathways (mesocorticolimbic DA systems), also affect brain stress pathways (CRF-HPA axis and the autonomic nervous system pathways) with increases in plasma adrenocorticotropic hormone (ACTH) and corticosterone, changes in heart rate and blood pressure and skin conductance responses. Regular and chronic use of these drugs is each associated with adaptations in these stress systems that are specific to each drug. For example, changes in heart rate and heart rate variability (HRV) are reported with regular and chronic alcohol use. Sustained increases in HPA axis function in the case of psychostimulants and tolerance to the inactivating effects of the drug

in the case of morphine, nicotine, and alcohol has also been demonstrated. These direct effects of drugs of abuse on major components of the physiological stress responses support their classification as "pharmacological" stressors.

Acute withdrawal states are associated with increases in CRF levels in cerebrospinal fluid, plasma ACTH, cortisol, and NE and EPI levels. Early abstinence is associated with high basal ACTH and cortisol responses, and a blunted or suppressed ACTH and cortisol response to pharmacological and psychological challenges in alcohol-dependent individuals and chronic smokers, while hyperresponsivity of HPA hormones in response to metyrapone has been reported in opiate and cocaine addicts. Furthermore, withdrawal and abstinence from chronic alcohol is also associated with altered sympathetic and parasympathetic responses, and altered noradrenergic responses to yohimbine challenge in early abstinence from cocaine have also been observed. The above changes highlight the significant effects of drug use and abuse on physiological stress responses.

Although acute administration of drugs increase mesolimbic dopamine, regular and chronic use of abusive drugs and acute withdrawal states decrease activity of the mesolimbic dopamine pathways with decreases in basal and stimulated dopamine reported in several preclinical studies. Basic science research shows chronic use of cocaine has also been shown to dramatically alter central noradrenergic pathways in the ventral and dorsal striatum, other areas of the forebrain, and the ventromedial prefrontal cortex. Human brain imaging studies corroborate these preclinical data, with reduced D2 receptors and dopamine transmission in the frontal and VS regions reported in alcohol-dependent individuals and cocaine-dependent individuals during acute withdrawal and protracted withdrawal (up to 3–4 months). Furthermore, blunted dopamine release in the VS and anterior caudate was also associated with choice to self-administer cocaine over money in human cocaine abusers. These changes are similar to the effects of prolonged and repeated stressors on mesolimbic dopamine and NE deficiency noted in the previous section, and suggest that chronic drug effects on extrahypothalamic CRF, noradrenergic or GC systems may directly impact the corticostriatal-limbic dopamine pathways.

On the other hand, acute, regular, and chronic exposure to drugs results in "sensitization" or enhanced behavioral and neurochemical response to drugs and to stress. Synaptic alterations in the VTA, NAc, and medial PFC modulated by glutamate effects on dopamine neurons, and CRF and noradrenergic effects on DA and non-DA pathways contribute to behavioral sensitization of stress and drugs of abuse. In addition, increased levels of brain-derived neurotrophic factor (BDNF) in the mesolimbic dopamine regions has been associated with increases in drug seeking during abstinence from chronic drug use. Furthermore, behavioral sensitization observed with drugs of abuse and with stress are associated with synaptic changes in mesolimbic dopamine regions, particularly the VTA, NAc, and amygdala, and such changes contribute to compulsive drug seeking. Thus, there are significant physiological, neurochemical, and behavioral alterations in stress and DA pathways associated with chronic drug use, which in turn could affect craving and compulsive seeking, maintenance of drug use, and relapse risk. It is not entirely clear how long these changes persist, and the extent to which there is recovery or normalization of these pathways and responses.

Altered Stress Responses and Increased Drug Craving with Chronic Drug Abuse

Clinical symptoms of irritability, anxiety, emotional distress, sleep problems, dysphoria, aggressive behaviors, and drug craving are common during early abstinence from alcohol, cocaine, opiates, nicotine, and marijuana. A mild "negative affect" and craving state ensues postwithdrawal associated with the alterations in stress and dopamine pathways. Severity of these symptoms have been associated with treatment outcomes with greater dependence and abstinence severity predictive of worse treatment outcomes.

Drug craving or "wanting" for drug is conceptually different from other anxiety and negative effect symptoms as it comes from desire or a wish for a rewarding or hedonic stimulus. However, with chronic drug use the terms craving and wanting often become associated with a physiological need, hunger and strong intent to seek out desired object, thereby representative of the more compulsive aspects of craving and drug seeking identified by addicted patients. In particular, craving and compulsive seeking is strongly manifested in the context of stress exposure, drug-related cues and drug itself and can become a potent trigger for relapse. Heightened craving or wanting of drug is associated with more severe drug use and is thought to represent the behavioral aspects of molecular and cellular changes in stress and dopamine pathways discussed in the previous section. Indeed some support for this idea comes from laboratory and imaging studies summarized below.

A series of studies in human addicted samples examined drug craving and stress responses in the context of stress and drug cues as compared to exposure to relaxing control cues. Significant increases in drug craving and subjective anxiety were observed in abstinent

addicted individuals compared to controls, but altered autonomic and HPA axis responses were observed with exposure to stress and nonstress drug cues as compared to neutral-relaxing cues in recently abstinent addicted individuals. For example, cocaine patients showed an enhanced sensitivity to emotional distress and physiological arousal and higher levels of drug craving to both stress and drug cue exposure compared to controls. Similarly, recovering alcohol-dependent individuals at 4-weeks abstinence show greater levels of basal heart rate and salivary cortisol levels compared to control drinkers. Upon stress and alcohol cue exposure, they showed persistently greater subjective distress, alcohol craving, and blood pressure responses, but suppressed heart rate and cortisol responses compared to controls. Interestingly, both cocaine-dependent and alcohol-dependent individuals show increased anxiety and negative emotions during drug cue exposure while social drinkers report lower levels of negative affect and anxiety with alcohol cue exposure. These data provide direct evidence of high drug craving and altered hedonic responses to both stress and drug cues in addicted individuals compared to social drinkers (see Fig. 23.2, reprinted with permission from Sinha, 2009). They also indicate that alteration in physiological stress responses are associated with high levels of stress-induced and cue-induced craving and distress states. The nature of the alterations are marked by increased emotional distress, heightened craving, altered basal responses, and blunted or suppressed physiological responses in abstinent addicted individuals compared to social drinkers.

Neural Correlates of Drug Craving in Addicted Samples

Many studies have also examined brain regions associated with craving in addicted individuals using functional magnetic resonance imaging procedures. Exposure to drug cues, known to increase craving, results in greater activation of the amygdala and regions of the frontal cortex, with gender differences in amygdala activity and frontal cortex response in cocaine-dependent individuals. Cue-induced craving for nicotine, methamphetamine, and opiates also activate regions of the prefrontal cortex, amygdala, hippocampus, insula, and VTA. Furthermore, exposure to stress cues increases brain response in paralimbic regions such as the anterior cingulate cortex, hippocampus and parahippocampal regions in healthy controls during stress while cocaine patients showed a striking absence of such activation. In contrast, patients had increased activity in the caudate and dorsal striatum region during stress, activation that was significantly associated with stress-induced cocaine craving ratings. Recent data from abstinent alcohol-dependent and cocaine-dependent patients have also shown higher activation in corticostriatal-limbic regions during relaxed trials, and particular blunted prefrontal and anterior cingulate responses during stress and drug cue exposure, and the latter being positive associated with drug craving responses.

Recent PET studies have also shown significant positive correlations between the dorsal striatum and drug cue-induced cocaine craving. These findings are

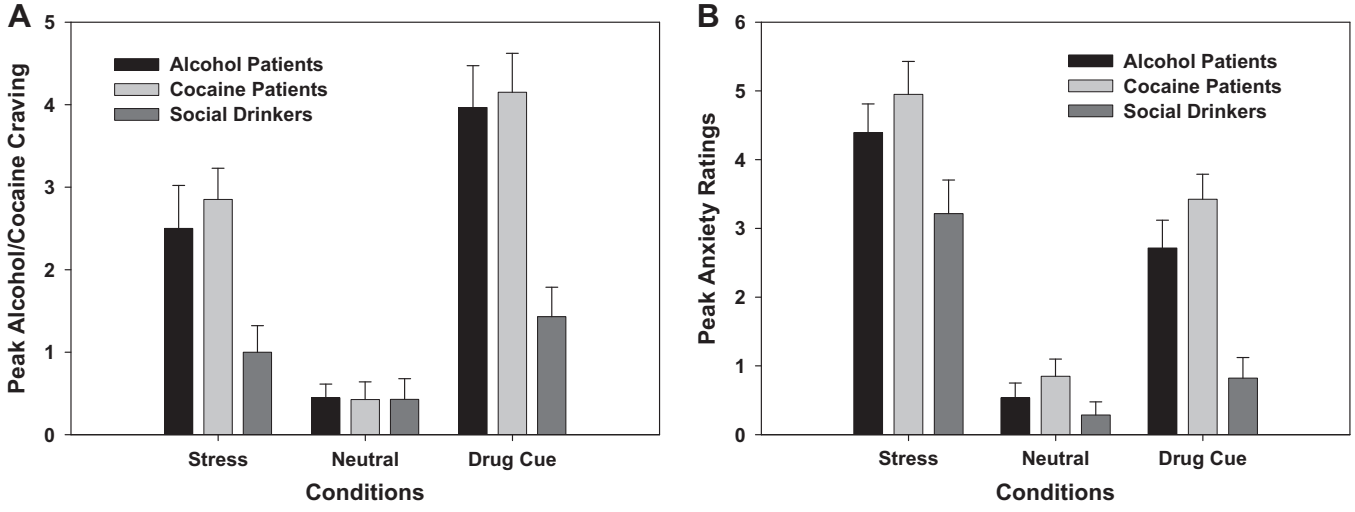


FIGURE 23.2 Mean and standard errors for peak craving and anxiety ratings during exposure to stress, drug cues, and neutral imagery conditions. **A:** Peak craving is significantly higher in abstinent alcohol-dependent and cocaine-dependent patients compared to social drinkers ($p < .0001$). **B:** Peak anxiety ratings is significantly higher in abstinent alcohol-dependent and cocaine-dependent patients compared to social drinkers ($p < .001$).

consistent with imaging studies with alcohol-dependent patients showing increased association between dorsal striatum regions and alcohol craving in response to presentation of alcohol-related stimuli. Using PET imaging with alcohol and cocaine patients, research has shown a significant association between dopamine D2 receptor binding in the VS and drug craving as well as motivation for self-administration. On the other hand, neuropsychological and imaging studies examining prefrontal executive functions, including impulse control, decision making, and set shifting, have shown executive function deficits and hypofrontal responses in addicted individuals compared to control volunteers. Together, these findings indicate that increased stress and cue-induced craving and compulsive drug seeking states in addicted individuals are associated with greater activity in the striatum, but decreased activity in specific regions of the cingulate and prefrontal cortex and related regions involved in controlling impulses and emotions.

Stress-induced Relapse

While several efficacious behavioral and pharmacological therapies in the treatment of addiction exist, it is well-known that relapse rates in addiction remain high. Exposure to stress, drug-related stimuli and drug itself each reinstate drug seeking behavior in animals and increase relapse susceptibility in addicted individuals. Such data underscore the need for specific attention to the chronic relapse susceptibility as a target in addiction treatment development.

In the last decade, a substantial number of preclinical studies have shown that brain CRF, noradrenergic and glutamatergic pathways contribute to reinstatement of drug seeking. Neuroadaptations associated with chronic drug use include overactive brain CRF and glutamatergic pathways, altered autonomic responses and underactive dopamine and GABA systems, and these changes may accompany the high craving states and relapse susceptibility associated with the chronic nature of addiction. Furthermore, using animal models of drug self-administration and relapse, preclinical studies have identified CRF antagonists, noradrenergic agents such as alpha-2-adrenergic agonists, alpha-1-adrenergic antagonists, and more recently glutamatergic agents, as playing a role in reducing stress-induced seeking in addicted laboratory animals.

Human research has also begun to identify markers of the stress and craving states that are predictive of relapse outcomes. Human laboratory models of drug craving and relapse risk have been developed to assess whether provoked drug craving and stress responses in the laboratory are predictive of relapse in the laboratory (e.g. smoking behavior) or relapse assessed prospectively following inpatient treatment. Thus, stress exposure increases drug craving in abstinent alcohol-, cocaine-, and nicotine-dependent individuals, and in each case stress-induced drug craving and dysregulated HPA axis responses are predictive of subsequent drug use and relapse. Thus, high levels of drug craving (stress or drug cue related) and dysfunctional HPA axis responses result in an enhanced susceptibility to addiction relapse.

NOVEL TREATMENT APPROACHES TO ADDRESS STRESS AND DRUG CRAVING

Findings from basic science and human laboratory and clinical outcome studies identify several pharmacological treatment targets to address stress-induced reinstatement of drug seeking and relapse susceptibility. Basic science data suggest CRF antagonists, noradrenergic agents, GABA and glumatergic agents could be promising in addressing stress-related relapse. Human laboratory studies using randomized, placebo controlled designs are now being conducted to screen these types of agents to assess their promise with regard to intermediate markers of stress-related relapse susceptibility. Such studies would target stress and cue-induced drug craving, craving-related anxiety, HPA measures and heart rate or HRV and also responses in specific brain regions. Early investigations have shown alpha-2-adrenergic agonists such as lofexidine and guanfacine to significantly decrease stress-induced and drug cue-induced drug craving and negative emotions in addicted samples. Recent evidence with Prazosin, an alpha-1-adrenergic antagonist has also shown prazosin-related reductions in stress-induced alcohol craving, anxiety, and negative emotions in alcohol-dependent individuals.

Behavioral interventions are also being developed and are in the process of validation to address the above outlined stress pathophysiology in addiction. Cognitive behavioral treatments that focus on drug craving and cognitive skill training in self-control are being considered. Stress interventions are being included in adapted cognitive behavioral interventions. A novel approach to redirect attention and address the cognitive bias that addicts show toward drugs and cravings has been developed and is currently being tested. Finally, mindfulness based relapse prevention adapted from mindfulness based stress reduction has been developed and is currently being tested to improve craving and stress dysfunction in addiction.

These new directions in pharmacological and behavioral interventions highlight the potential of developing

new therapies to address stress-related pathophysiology in addiction.

CONCLUSIONS AND FUTURE DIRECTIONS

This chapter focuses on the accumulating evidence from preclinical, clinical, and population studies that highly stressful situations and chronic stress increases addiction vulnerability, which increases both the risk of developing addiction and risk of relapse. The types of stressors that increase addiction risk are identified in Table 23.1. The stressors tend to be highly emotionally distressing events that would be uncontrollable and unpredictable for both children and adults. The themes range from loss, violence and aggression, poor support, interpersonal conflict, and isolation and trauma. There is also evidence for a dose-dependent relationship between accumulated adversity and addiction risk, with the greater the number of stressors an individual is exposed to, the higher the risk of developing substance use disorders. Work-related stressors have weaker support but individual-level variables such as trait negative emotionality and poor self-control (possibly similar to poor executive function) appear to also contribute uniquely to addiction risk. Exposure to such stressors early in life and accumulation of stress (chronicity) result in neuroendocrine, physiological, behavioral, and subjective changes that tend to be long lasting and adversely affect development of brain systems involved in learning and motivation and stress-related adaptive behaviors. Research that directly addresses stress-related neurobiological changes and its association with behavioral outcomes is sorely needed. Evidence to clarify the contribution of stress on alterations in mesolimbic dopamine activity and it's association to drug use is also needed.

A review of evidence indicating the effects of drug use and abuse on stress responses and dopamine transmission is presented. Altered emotional and motivational responses that are associated with craving and relapse to drug use are reviewed. While substance abuse results in changes in stress and DA pathways involved in motivation, self-control and adaptive processes necessary for survival, whether such changes enhance drug seeking or craving and drug use behaviors is lacking. For example, studies on whether prior exposure to licit and illicit drugs modifies the association between stress and drug self-administration are rare. While there are specific neuroadaptations in reward and associated regions, it is also important to examine which of these changes are involved in increasing drug intake and supportive of addictive processes such as progressive loss of control, persistence of craving and escalating drug self-administration. As stress also increases risk of mood and anxiety disorders which are highly comorbid with addiction, it is important to examine whether there are specific stress-related factors that contribute to risk for mood and anxiety disorders and addiction risk. Exploration of gene–environment interactions could be particularly helpful in answering such questions.

A review of recent studies on stress-induced reinstatement to drug seeking, drug craving, and relapse susceptibility is also provided. Clinical implications include the development of new assessment procedures and markers that will be useful in identifying those who are at particular risk for stress-related relapse and testing of novel pharmacological therapies that target the link between stress and relapse risk. As shown in Fig. 23.2, addicted individuals show enhanced sensitivity to craving and greater anxiety in stress and drug-related situations, but whether such altered responses represent transitions due to chronic drug use or chronic stress states needs to be further examined. Research on the mechanisms by which chronic stress and drug use alter executive functions that are involved in adaptive behavioral responses is needed. Efficacious behavioral treatments focus on improving coping response. However, stress exposure and chronic distress decrease stress adaptive and coping mechanisms, and hence treatments that focus on enhancing coping may not be suitable for those with stress-related addiction risk. Development of new interventions that target self-control, especially in the context of stress is needed. Systematic research on these questions will lead to a greater understanding of how stress is associated with relapse. Furthermore, such research may be significant in developing new treatments targets to reduce relapse, risk in addictive disorders.

SEE ALSO

The Intergenerational Transference of Addiction, Relapse and Lapse, Adolescent Substance Use: Symptoms and Course, Implicit and Associative Processes in Addiction, Craving and Expectancies

List of Abbreviations

ACTH	adrenocorticotropic hormone
BDNF	brain-derived neurotrophic factor
CRF	corticotrophin releasing factor
DA	dopaminergic
EPI	epinephrine
HPA axis	hypothalamic-pituitary-adrenal
HRV	heart rate variability
mPFC	medial prefrontal cortex
NAc	nucleus accumbens
NE	norepinephrine
PET	positive emission tomography

PVN paraventricular nucleus
VS ventral striatum
VTA ventral tegmental area

Glossary

Catecholamines amines that act as a hormone or a neurotransmitter. Includes epinephrine (adrenaline), norepinephrine (noradrenaline), and dopamine.
Corticosterone a steroid hormone produced in the adrenal cortex that affects carbohydrate, potassium, and sodium metabolism and plays a key role in stress experiences and regulation.
Corticotrophin releasing factor (CRF) a hypothalamic neuropeptide thought to mediate autonomic and behavioral responses to stress.
GABA gamma-aminobutyric acid is the primary inhibitory neurotransmitter of the central nervous system.
Glucocorticoids (GC) a class of steroid hormones involved predominately in the metabolism of carbohydrates, that exert an anti-inflammatory effect and include cortisol (hydrocortisol).
Hypothalamic-pituitary-adrenal (HPA) axis the interaction between the hypothalamus, pituitary gland, and the adrenal glands, thought to regulate many processes, including our reaction to stress.
LC locus coeruleus
Noradrenergic areas of the body that are affected by or produce norepinephrine.
Nucleus accumbens (NAc) a group of neurons located beneath the frontal lobe, forming the main part of the ventrial striatum. It plays a role in the reward response and in experience of pleasure.
Prefrontal cortex (PFC) located in the anterior region of the frontal lobes, and thought to be responsible for higher order executive functions.
Sympathoadrenal medulary system activated when exposed to sudden stress. Results in the stimulation of adrenergic neurons of the hypothalamus, then the release of epinephrine from the adrenal medulla and norepinephrine from the sympathetic ganglia.
Ventral tegmental area (VTA) located in the midbrain and containing dopamine cell bodies. The mesolimbic dopamine pathway originates here.

Further Reading

Adinoff, B., Stein, E.A. (Eds.), 2011. Neuroimaging in Addiction, first ed. Wiley-Blackwell, John Wiley & Sons Publication, West Sussex, UK.
Arnsten, A.F., 2009. Stress signalling pathways that impair prefrontal cortex structure and function. Nature Reviews Neuroscience 10 (6), 410–422. Review.
Arnsten, A., Mazure, C.M., Sinha, R., 2012. This is your brain in meltdown. Scientific American Apr;306 (4), 48–53.
Epstein, D.H., Willner-Reid, J., Vahabzadeh, M., Mezghanni, M., Lin, J.L., Preston, K.L., 2009. Real-time electronic diary reports of cue exposure and mood in the hours before cocaine and heroin craving and use. Archives of General Psychiatry 66 (1), 88–94.
Goldstein, R.Z., Volkow, N.D., 2011. Dysfunction of the prefrontal cortex in addiction: neuroimaging findings and clinical implications. Nature Reviews Neuroscience 12 (11), 652–669.
Gordon H.W., Majewska M.D., and Thadani P.V. (Eds.) (2002) Special Issue: Stress and Drug Abuse. Psychoneuroendocrinology, 27(1/2).
Kabat-Zinn, J., 1989. Full Catastrophe Living: Using the Wisdom of Your Body and Mind to Face Stress, Pain and Illness. Piatkus, London.
Koob, G., Kreek, M.J., 2007. Stress, dysregulation of drug reward pathways, and the transition to drug dependence. American Journal of Psychiatry 164 (8), 1149–1159. Review.
McEwen, B.S., 2007. Physiology and neurobiology of stress and adaptation: central role of the brain. Physiological Reviews 87 (3), 873–904. Review.
McEwen, B.S., Lasley, E.N., 2004. The End of Stress as We Know It. National Academies Press.
Steckler, T., Kalin, N., Reul, J.M.H.M. (Eds.), 2005. Handbook on Stress and the Brain: Part 2: Stress: Integrative and Clinical Aspects. Elsevier Science, Amsterdam.
Sinha, R., 2008. Chronic stress, drug use and vulnerability to addiction. Annals of the New York Academy of Sciences: Addiction Reviews 1141, 105–130.
Sinha, R., Shaham, Y., Heilig, M., 2011. Translational and reverse translational research on the role of stress in drug craving and relapse. Psychopharmacology (Berl) 218 (1), 69–82. Epub 2011 Apr 15.
Sinha, R., 2011. New findings on biological factors predicting addiction relapse vulnerability. Current Psychiatry Reports 13 (5), 398–405. Review.

Relevant Websites

http://yalestress.org/podcasts.html – The Yale Stress Center.
http://www.cpbn.org/program/%5Bfield_episode_cpbi_program%5D/episode/coping-stress – Connecticut Public Broadcasting Network's Radio.
http://www.apa.org/news/press/releases/stress/index.aspx – The American Psychological Association.
http://www.drugfree.org/join-together – The Partnership at Drugfree.org.
http://archives.drugabuse.gov/NIDA_Notes/NNVol14N1/Stress.html – Studies Link Stress and Drug Addiction.
http://www.hbo.com/addiction/understanding_addiction/141_stress.html – HBO: Addiction: Understanding Addiction.

CHAPTER

24

Self-Medication

Virginia Gil-Rivas, Linda McWhorter
University of North Carolina, Charlotte, Charlotte, NC, USA

OUTLINE

Substance use, both licit (e.g. alcohol, nicotine) and illicit (e.g. marijuana, cocaine, and heroin), can have serious health, social, legal, and financial consequences for individuals and society. The estimated combined annual cost of the medical, economic, criminal, and social consequences of substance use in the United States is approximately half a trillion dollars.

In 2008, approximately 8% of the US population aged 12 years or older reported using illicit substances during the past month. Further, 5% of adults reported heavy drinking (two or more drinks per day) and 15% reported binge drinking (five or more drinks in one occasion). Of those using drugs, it was estimated that 8.9% met diagnostic criteria for a substance use disorder (SUD) (i.e. substance use dependence or abuse) as defined by the *Diagnostic and statistical manual of mental disorders, 4th edition* (DSM-IV). SUDs are characterized by (1) a compulsion to seek and take a particular psychoactive substance; (2) lack of control over substance intake; (3) the presence of negative affective states (i.e. anxiety, dysphoria, and irritability) and somatic symptoms associated with discontinued use; and (4) a tendency to relapse after a period of abstinence.

Researchers and clinicians remain puzzled by the fact that individuals continue to use substances despite negative consequences. Multiple explanations have been offered for substance use and addiction, and perhaps one of the most influential explanations is the self-medication hypothesis (SMH) proposed by Khantzian. The SMH views substance use as an effort to cope with stress and distressing emotions (e.g. depression, anxiety, anger) among individuals with deficits in emotion regulation, self-care, interpersonal skills, and self-esteem that may or may not be related to a psychiatric disorder. In other words, SUDs are disorders characterized by self-regulation deficits. Further, the theory also suggests that individuals select specific substances based on their expected ability to reduce or increase specific emotions. For example, opiate use is thought to be related to efforts to manage hostile and violent emotions, while alcohol is frequently used as a way of facilitating social interaction and the expression of emotions among individuals with deficits in these areas.

In the short-term, the use of substances to manage distressing emotions can be adaptive as these substances provide individuals with a reprieve from intense emotional states. In the longer term, chronic use can contribute to the development of SUDs. Once individuals transition from occasional use to addiction, substance use withdrawal leads to distressing emotions

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00024-3

and somatic symptoms that contribute to relapse. Recent developments on the neurobiology of addiction provide some support for the idea that substance use has the potential of eliciting intensive positive affective states early in the process of addiction. Also, these findings suggest that in the longer term chronic substance use leads to neurobiological changes that are associated with psychiatric symptomatology.

NEUROBIOLOGICAL BASIS OF SUBSTANCE USE AND ADDICTION

The Brain Reward Circuit

The area of the brain known as the "reward circuit" or pleasure pathways that is located in the limbic regions including the ventral tegmental and nucleus accumbens has been shown to be activated at varying degrees by both licit and illicit substances. Although different drugs may affect these pathways somewhat differently, they all lead to the release of dopamine, a neurotransmitter that is associated with the experience of reward or pleasure. Importantly, the amount of dopamine released in the brain as a result of substance use is 2–10 times greater than the release associated with natural stimuli (e.g. food). The amount of dopamine released and the pattern of activation in the reward pathways explains the experience of euphoria or "high" associated with a particular substance. Thus congruent with the SMH, in the initial stages of addiction, substance use may be viewed by individuals as a way of achieving a reprieve from negative emotional states. After a period of experimentation, individuals may select a substance or family of substances (e.g. stimulants, opiates) that are the most effective in producing the desired affective state. It is important to consider that individuals' initial motivation for substance use may or may not reflect a desire for ameliorating negative affective states. It is also possible that individuals' motivation to use reflects a desire for the "high" associated with a substance, a need to regulate their overall emotional experiences, or an effort to obtain desired social rewards.

Research on the neurobiological basis of addiction has shown that dopamine not only acts as a "reward" but also plays a role in the encoding reward expectancies and of memories of events or stimuli that are associated with its release. As a result, simply the expectation of the reward leads to dopamine release in the reward pathways. In addition, stimuli associated with substance use (e.g. specific settings, emotional states, and images) may elicit dopamine release and increase drug craving and substance seeking behaviors. In other words, learning processes may play a role in the maintenance of substance use.

Over time, chronic substance use leads to dysregulation or a state of hedonic allostasis in the reward pathways. Further, these changes persist long after discontinuing substance use. In particular, chronic use is associated with reductions in basal dopamine levels, the amount of stimulated dopamine released, and in the number of dopamine receptors in these areas of the brain. Downregulation of positive reward systems and the negative affective states (i.e. anxiety and dysphoria) that accompany substance use withdrawal seem to underlie continued use despite negative health and social consequences. Thus, these neurological changes seem to explain the transition from use that is motivated primarily by the desire to obtain the "high" or the desired affective states associated with a particular substance (positive reward), to use as a way of reducing negative affective states (negative reward) when the substance is not present. In other words, addicted individuals use substances with the goal of gaining a sense of "normalcy" rather than to simply achieve a high.

Stress and Substance Use

In addition to the pleasure pathways, changes in the area of the brain associated with stress regulation play a role in substance use initiation, maintenance, and relapse. The stress response involves the activation of the hypothalamic-pituitary-adrenal (HPA) axis and the sympathetic adrenomedullary (SAM) system in response to stressful stimuli. The HPA activation begins with the release of corticotropin-releasing factor (CRF) from the paraventricular neurons of the hypothalamus. In turn, CRF stimulates the synthesis and release of adrenocorticotropin hormone (ACTH) by the anterior pituitary. ACTH stimulates the adrenal cortex to synthesize and release glucocorticoids (cortisol in humans). Activation of the SAM results in the release of catecholamines (adrenaline and noradrenaline). The activation of both the SAM and HPA results in the physiological changes (i.e. increase heart rate, blood pressure, respiration, and circulating glucose) associated with the physiological stress response. Importantly, the stress response also involves an increase of CRF in the amygdala, an area of the brain that plays an important role in anxiety-related affective states.

Available research suggests that early in the addiction process the activation of the HPA by both stress and substance use contributes to the sensitization of the reward pathways by glucocorticoids. In fact, animal studies have shown that glucocorticoids will trigger the release of dopamine in these pathways. Further, glucocorticoids augment the positive reinforcing effects of substances by altering the dopamine signaling in these areas. As individuals continue to use substances,

stress and glucocorticoids contribute to dysregulation of the brain reward pathways leading to increased substance use. Once individuals become addicted, high levels of glucocorticoids and other substances associated with the stress response lead to internally generated negative affective states. Additionally, substance use withdrawal is associated with increases in anxiety-like affective states. These negative affective states seem to contribute to craving and substance-seeking behaviors and in turn to continued substance use. In sum, the combined effects of chronic stress and substance use can result in neuroendocrine changes that can lead to dysfunction in the brain reward pathways and contribute to the development of addiction.

Substance use, however, does not always lead to the development of SUDs; rather it seems that genetic, individual, social, and environmental factors make some individuals more vulnerable to transition from occasional use to addiction. At the individual level, neurobiological studies suggest that differences in HPA axis activity may help explain why some individuals are more vulnerable to the development of SUDs compared to others. For example, adults who secrete high levels of cortisol are also high dopamine releasers and report greater effects from psychostimulants (e.g. cocaine, methamphetamines) compared to individuals with low cortisol and low dopamine levels. In addition, exposure to traumatic events (e.g. combat, physical assault, major disasters) and chronic stress have been shown to be associated with increases in substance use. The clinical and developmental literatures also suggest that exposure to unresponsive and insensitive caregiving environments and trauma (e.g. physical and sexual abuse) early in life affects the developing brain and the physiological systems associated with the stress response. Further, individuals with histories of exposure to adverse childhood environments seem to have diminished capacity to regulate their negative emotions and behavior, and to cope effectively with stress. As suggested by the SMH, these deficits in regulatory capacity could explain the increased risk for the development of SUDs later on in life among these individuals. Chronic stress and traumatic experiences in both childhood and adulthood seem to contribute not only to the development of SUDs, but also to the emergence of some psychiatric disorders. In fact, some researchers have characterized psychiatric disorders as chronic stress states.

SUBSTANCE USE AND PSYCHIATRIC DISORDERS

Current estimates among community samples indicate that 45–72% of adults with SUDs also have at least one co-occurring psychiatric disorder. Among individuals diagnosed with a psychiatric disorder (e.g. schizophrenia, major depression, posttraumatic stress disorder (PTSD), attention deficit disorder) approximately 30% also meet criteria for a concurrent SUD. Importantly, the presence of co-occurring disorders has been found to compromise both mental health and substance abuse treatment outcomes. Neurobiological studies have identified some abnormalities in the stress pathways that are involved in both psychiatric disorders and SUDs. In particular, researchers have focused on a potential common neurobiological basis for SUDs, depression/mood disorders, anxiety disorders, and schizophrenia.

Mood Disorders

Substantial symptom overlap between mood disorders (e.g. major depression, dysthamia, bipolar disorder) and SUDs exists. For example, irritability, sleep difficulties, anxiety, and concentration difficulties are frequently reported among individuals with major depression and those undergoing substance use withdrawal. Large population studies have found that about 32% of individuals with any mood disorder also met DSM-IV criteria for an SUD. At the neuroendocrine level, individuals diagnosed with major depression present abnormalities in the reward and stress pathways that are also seen among individuals with SUDs. For example, in response to CRF stimulation, depressed individuals have been found to have lower levels of ACTH and cortisol responses compared to healthy individuals, suggesting increased HPA axis activity. Longitudinal studies examining the temporal association between SUDs and mood disorders generally indicate that mood disorders emerge after the onset of SUDs. These findings suggest that neurobiological changes associated with addiction may contribute to the emergence of mood disorders, and thus contrary to the SMH, depressed mood seems to be a consequence of substance use.

Anxiety Disorders

Epidemiological studies have found that individuals with an anxiety disorder are more likely to also have SUD compared to those without such histories. Mixed empirical findings have emerged regarding the temporal association between SUDs and anxiety disorders. Some studies have found that the presence of a lifetime anxiety disorder (e.g. general anxiety, social phobia) is associated with an increased risk for the development of SUDs later on in life. In contrast, a recent study found that alcohol dependence preceded the onset

of several anxiety disorders (i.e. general anxiety, panic) but followed the onset of other disorders (i.e. specific phobia and social phobia). Congruent with the SMH, these findings suggest that some individuals may have started using substances as a way of managing anxiety-related symptoms. In support of this idea, some studies have found that individuals with lower tolerance for internally generated negative affective states and who rely on emotional avoidance are more likely to have co-occurring anxiety and SUDs. Contrary to the SMH, however, it seems that in some cases substance use may have contributed to the development of some anxiety disorders.

PTSD is an anxiety disorder that seems to be strongly related to SUDs. Specifically, current estimates indicate that among individuals seeking substance abuse treatment, 19–62% also meet diagnostic criteria for PTSD. Rates of co-occurring PTSD in community samples of adults range from 52 to 35% for men and 28 to 27% for women. Investigators have identified potential pathways that may link these disorders. Specifically, the following pathways have been suggested: (1) SUDs can increase the likelihood of exposure to trauma and thus contribute to the development of PTSD; (2) chronic substance use and the anxiety-like symptoms associated with withdrawal from use may make individuals with SUDs more vulnerable to the development of PTSD after trauma exposure; (3) consistent with the SMH, individuals may use substances in an attempt to alleviate memories and painful symptoms associated with PTSD; and (4) contrary to the SMH, substance use may interfere with the processing of traumatic memories and thus contribute to the exacerbation of PTSD symptoms.

Epidemiological studies have shown that typically PTSD precedes the development of SUDs providing support for the component of the SMH suggesting that individuals use substances with the goal of alleviating negative affective states. Neurobiological and neuroimaging studies suggest some common pathways that may help explain their co-occurrence. Specifically, the HPA axis and SAM are involved in stress, PTSD, and SUDs. Recent studies have identified high levels of CRF in the amygdala and chronically increased levels of catecholamines among individuals with PTSD which are suggestive of chronically increased stress responses and are believed to contribute to an increased risk of developing SUDs.

Schizophrenia-Spectrum Disorders

Schizophrenia-spectrum (i.e. schizophrenia, schizophreniform, or schizoaffective) disorders show high levels of co-occurrence with SUDs. Epidemiological estimates indicate that approximately 50% of individuals with these disorders have a lifetime history of substance use and 27% meet criteria for an SUD. These rates are significantly higher than those among the general population. In particular, alcohol and marijuana use are common among individuals diagnosed with schizophrenia. Researchers have suggested that SUD and schizophrenia may share some common biological, individual, and social risk factors that may explain their high rates of co-occurrence. Of particular relevance, are changes in the reward pathways involving dopamine and glutamate that are associated with schizophrenia. These changes may render individuals more vulnerable to SUDs. In addition, the SMH has been a dominant framework for understanding their co-occurrence as individuals report using substances as a way of reducing symptoms such as social withdrawal, apathy, cognitive impairment, and as a way of managing the side effects of antipsychotic medications.

SUPPORT FOR THE SMH

Substance Use to Manage Negative Affective States

The SMH has garnered a great deal of interest among clinicians and researchers attempting to find potential explanations for SUDs. Despite the intuitive appeal of this hypothesis, the empirical literature provides mixed support for each of its components. The first component of the SMH, that is, the assumption that substance use represents a way of managing negative affective states or symptoms among individuals with self-regulation deficits, has received the most attention in the literature. Studies with nonclinical samples of adolescents and young adults suggest that self-medication may not be the only or the most important factor predicting the initiation of substance use in this group. In particular, it seems that individual differences in sensation seeking and the social meaning given to substance use play an important role in the initiation of substance use. Compelling evidence for the limited role of management of negative mood in explaining substance use among adolescents and young adults has been provided by studies using momentary assessment methodology. This methodology allows for the examination of the temporal association between negative mood states and substance use. For example, adolescents high in novelty seeking have been found to be more likely to use alcohol and marijuana when they are experiencing high levels of positive mood, rather than when experiencing distressing emotions. In contrast, a study of adults that also made use of momentary assessment methodology showed that both anxious mood and pleasant mood states were associated with increases in later alcohol consumption.

Further, among those reporting anxious mood, alcohol consumption contributed to reductions in anxiety. Studies using more traditional methodologies suggest that compared to emotional states, personality characteristics are stronger predictors of substance use among youths. Specifically, trait anxiety, trait aggression, and attention and behavioral difficulties in school have been found to be associated with an increased risk for developing SUDs later in life accounting for risk factors such as parental education and mental health. Overall, these findings suggest that individuals may have a variety of reasons for using substances and that the management of negative affective states may be the primary reason for only a subset of individuals.

The evidence supporting the SMH among individuals with SUDs and those with co-occurring psychiatric conditions is also mixed. Substance use motivated by the need to alleviate negative affective states seems to be common among individuals with SUDs and those with psychiatric disorders; however, this is not the only reported reason for use. A review of the literature of self-reported reasons for substance use among individuals with schizophrenia-spectrum disorders indicated that while 2–86% of respondents endorsed substance use as a way of relieving negative affective states, 62–95% of these individuals reported using to increase pleasure and 35–98% reported using to achieve the "high" or intoxicating effects.

It is important to consider that although individuals may use substances in an attempt to relieve negative or distressing emotions, substance use may also elicit feelings of shame, guilt, and powerlessness leading to an exacerbation of these negative states. In light of these findings, some addiction models suggest that negative affective states and substance use form a feedback loop in which changes in one of these factors influence changes in the other. Support for this feedback loop was reported in a recent longitudinal study that found that changes in negative affect (i.e. anger and depression) were associated with alcohol use and that changes in drinking were associated with changes in negative affect after alcohol treatment.

The SMH has also been used to understand substance use relapse after a period of abstinence. The literature suggests that both intrapersonal and interpersonal factors are common precursors to substance use relapse. Intrapersonal factors are those generated by the individual and include negative emotions (i.e. depressed and anxious mood), negative physiological states (e.g. pain), positive emotional states, testing personal control, and urges and temptations. In contrast, interpersonal factors involve external or environmental influences such as social pressure, and interpersonal conflict. Consistent with the SMH, anger, frustration, interpersonal conflict, and depressed mood have been shown to be the most commonly reported antecedents of substance use relapse. Similarly, among individuals with co-occurring SUDs and psychiatric disorders, depression, boredom, insecurity, anxiety, loneliness, and sleep difficulties frequently precede relapse to substance use after treatment.

Among individuals with psychiatric disorders, substance use early in the addiction process may serve to reduce psychiatric symptoms; however, over time substance use may lead to an exacerbation of symptoms. In fact, some investigators have proposed that substance use among addicted individuals may have a rebound effect, leading to the exacerbation of psychiatric symptoms rather than symptom improvement. From this perspective, the neurobiological changes associated with chronic substance use and the ability of these substances to occasionally elicit symptoms improvement may help explain the maintenance of this behavior. Further, some substances may help alleviate some symptoms and contribute to an increase of other symptoms on termination of their psychoactive effects. Congruent with the proposed rebound effects, recent empirical findings suggest that although increases in mental health symptoms (i.e. anxiety, depression, and sleep disturbances) may be reported before substance use relapse, substance use does not seem to ameliorate these symptoms. For example, a recent study tested the SMH and the "rebound hypothesis" among adults with SUDs and those with co-occurring disorders. Congruent with the SMH, individuals reported increases in symptoms before substance use relapse. In contrast, a majority (about 60%) of participants demonstrated an exacerbation of symptoms immediately after use and 2 weeks later, and only 25% of the sample reported any symptom improvement after use. Further, substance use resulted in an exacerbation of symptoms that were related to individuals' lifetime mental health diagnosis. Specifically, individuals with bipolar disorders experienced an exacerbation of depressive symptoms and those diagnosed with schizophrenia reported increases in psychotic symptoms. Similarly, a study with adolescents with SUDs found that psychiatric and depressive symptoms preceded substance use relapse; however, substance use contributed to increased symptom severity for over half of these individuals, while only 20% reported symptom improvement. Overall, these findings provide support for the rebound effects of substance use among individuals with SUDs and co-occurring disorders and mixed support for the SMH.

Substance Specificity

By and large, support for the "specificity" component of the SMH suggesting that individuals select specific substances based on their expected pharmacological effects is based on case studies and anecdotes from

substance-abusing patients receiving psychotherapy. The empirical literature, however, provides mixed support for this assertion. Importantly, the available literature testing this component of the SHM has focused on individuals who were receiving substance abuse treatment, and thus it is unclear to what extent early in the addiction process specific substances were selected by individuals based on their ability to induce the desired effects.

Studies among individuals with psychiatric disorders have not found strong evidence for a match between the symptoms that characterize the disorder and the type of substance used by individuals. Partial support for the "specificity" component of the SMH emerged in a study of individuals receiving inpatient mental health treatment. Specifically, participants with depressive disorders were more likely to meet criteria for opiate abuse compared to those diagnosed with a bipolar disorder or schizophrenia. Further, individuals with schizophrenia were more likely to abuse alcohol compared to the other two groups, while individuals with bipolar disorder were more likely to meet criteria for marijuana abuse. However, it is important to note that in this study a majority of participants reported using more than one substance. Other studies have not found support for the specificity component of the SMH among individuals with SUDs and those with co-occurring disorders.

SUMMARY

The SMH suggests that (1) individuals with self-regulation deficits use substances in an attempt to manage negative or distressing affective states and (2) individuals select specific substances based on their ability to elicit desired affective states.

Congruent with the SMH, substance use early in the addiction process can successfully lead to reductions in negative affect among individuals who lack more adaptive emotion and self-regulation skills. In the longer term, chronic substance use contributes to neurobiological changes in the reward and stress pathways, which play an important role in motivation, learning, and behavior. These changes may underlie the transition from substance use motivated by achieving the "high" or relief associated with substance use to use that is motivated by the need to regain a sense of normalcy.

The empirical literature supporting the SMH is mixed. Among nonclinical samples of adolescents and young adults, the primary motivation for substance use does not seem to be the management of negative affective states. Rather, sensation seeking and the social meaning given to the use of these substances seem to be the primary motivating force for substance use. Importantly, positive rather than negative affective states seem to increase the likelihood of substance use among adolescents.

Among adults, a majority of studies aiming to test the SMH have been conducted with clinical samples of individuals diagnosed with SUDs and/or psychiatric disorders. Available research indicates that these individuals report using to relieve negative affective states as well as to obtain the "high" or pleasurable effects associated with a particular substance. Further, the evidence suggests that substance use increases rather than alleviates negative affective states. It is possible that during the intoxication period substance use provides relief but leads to increases in symptomatology once the effects have worn off. In fact, these findings provide support for the "rebound" hypothesis and addiction models that propose a feedback loop between substance use and negative affective states.

Support for the second component of the SMH, that is, the idea that individuals select "specific" substances based on their effects on relieving distress or negative affective states, is rather limited. Specifically, the available studies have not examined whether individuals experiment with different substances early in the addiction process until they identify the substance that most effectively elicits the desired effects. Studies among individuals diagnosed with SUDs and a psychiatric disorder also provide mixed support for the "specificity" assumption. These individuals typically report using multiple substances. Further, a pattern of substance selectivity based on symptoms or psychiatric diagnosis has not been consistently identified.

In conclusion, empirical support for the SMH is mixed. Reliance on cross-sectional studies and methodologies that do not allow for an adequate assessment of the temporal relationship between negative affective states and substance use has not allowed for an adequate test of the SMH at all stages of the addiction process. Future research is needed to clarify the role of negative affective states in the initiation, progression, and maintenance of substance use.

SEE ALSO

Behavioral Economic Factors in Addictive Processes, The Biopsychosocial Model of Addiction, Cognitive Factors in Addictive Processes, Emotions and Addictive Processes, Contextual Factors in Addiction, Relapse and Lapse, Stress and Addiction, Neural Correlates of Craving for Psychoactive Drugs, Models of Relationships between Substance Use and Mental Disorders, Substance Use and Mood Disorders, Substance Use in Response to Anxiety Disorders

Glossary

ACTH adrenocorticotropin hormone – A hormone produced by the pituitary in response to stimulation by CRF. This hormone stimulates the cortex of the adrenal glands to release glucocorticoids as part of the physiological response to stress.

CRF corticotropin-releasing factor – a hormone released by the hypothalamus as part of the stress response. This hormone stimulates the pituitary to produce ACTH.

DSM-IV *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* – a manual published by the American Psychiatric Association that provides standard criteria for the classification of mental disorders.

HPA hypothalamic-pituitary-adrenal axis – a component of the neuroendocrine system that plays an important role in the body's physiological response to stress. Activation of this system leads to the release of glucocorticoids by the adrenal glands.

PTSD posttraumatic stress disorder – an anxiety disorder that can develop after exposure to traumatic events. This disorder is characterized by persistent thoughts or memories about the traumatic event, sleep difficulties, emotional numbness, and high levels of emotional arousal.

SAM sympathetic adrenomedullary system – a neuroendocrine system involved in the body's physiological response to stress. Activation of this system during stress leads to the release of catecholamines (epinephrine and norepinephrine).

SMH self-medication hypothesis – a hypothesis that suggests that addiction is the result of self-regulation deficits. Specifically it proposes that (1) individuals use substances as a way of coping or managing with negative emotional states and (2) individuals select specific substances for their expected ability to reduce or increase specific emotions.

SUD substance use disorder – a disorder that involves the repeated use or abuse of psychoactive substances despite negative consequences.

Further Reading

Blume, A.W., Schmaling, K.B., Marlatt, G.A., 2000. Revisiting the self-medication hypothesis from a behavioral perspective. Cognitive and Behavioral Practice 7, 379–384.

Brady, K.T., Sinha, R., 2005. Co-occurring mental and substance use disorders: the neurobiological effects of chronic stress. American Journal of Psychiatry 162, 1483–1493.

Cicchetti, D., Curtis, W.J., 2006. The developing brain and neural plasticity: implications for normality, psychopathology, and resilience. In: Cicchetti, D., Cohen, D. (Eds.), Developmental Psychopathology 2: Developmental Neuroscience, second ed.). Wiley, New York, pp. 1–64.

Gregg, L., Barrowclough, C., Haddock, G., 2007. Reasons for increased substance use in psychosis. Clinical Psychology Review 27, 494–510.

Khantzian, E.J., 1997. The self-medication hypothesis of substance use disorders: a consideration and recent applications. Harvard Review of Psychiatry 4, 231–244.

Khantzian, E.J., 2003. Understanding addictive vulnerability: an evolving psychodynamic perspective. Neuro-Psychoanalysis 5, 5–21.

Koob, G.F., Le Moal, M., 2008. Addiction and the brain antireward system. Annual Review of Psychology 58, 29–53.

McCarthy, D.M., Tomlinson, K.L., Anderson, K.G., Marlatt, G.A., Brown, S.A., 2005. Relapse in alcohol- and drug-disordered adolescents with comorbid psychopathology: changes in psychiatric symptoms. Psychology of Addictive Behaviors 19, 28–34.

Tomlinson, K.L., Tate, S.R., Anderson, K.G., McCarthy, D.M., Brown, S.A., 2006. An examination of self-medication before and after alcohol or drug relapse. Addictive Behaviors 31, 461–474.

Uhart, M., Wand, G.S., 2008. Stress, alcohol and drug interaction: an update of human research. Addiction Biology 14, 43–64.

Volkow, N.D., Fowler, J.S., Wang, G.J., Baler, R., Telang, F., 2009. Imaging dopamine's role in drug abuse and addiction. Neuropharmacology 56, 3–8.

Witkiewitz, K., Villarroel, N.A., 2009. Dynamic association between negative affect and alcohol lapses following alcohol treatment. Journal of Consulting and Clinical Psychology 77, 633–644.

Relevant Websites

http://drugabuse.gov/nidahome.html – National Institute on Drug Abuse.

http://www.samhsa.gov/ – Substance Abuse and Mental Health Services Administration.

CHAPTER

25

Contextual Factors in Addiction

John M. Roll

Washington State University, Spokane, WA, USA

OUTLINE

INTRODUCTION

Like all organisms, humans exist in a complex environment. In this chapter, it is argued that our behavior is largely controlled by our perceptions of this environment. It is useful to consider how this environment can be parsed. Perhaps the simplest conceptualization postulates that the environment consists of both external and internal components with respect to an individual – a perceiver. The external components are those things we typically consider to be environmental. The external environment includes things that the perceiver senses as being outside their anatomy such as air purity, ambient temperature, and elevation. The internal environment consists of factors that are inside the perceiver. This includes such observable states as hunger, thirst, fatigue, illness, and various affective states. The constant interplay between the external and the internal environments provide a molar framework in which to explain behavior. For instance, a person standing outside may don a coat as a blizzard approaches because the ambient temperature has dropped and made increased demands on their metabolic system to keep them warm. However, a person who has a high fever may welcome the approach of cold weather and not don a coat because they anticipate the relief from their feverish state.

However, the richness of human behavior cannot be fully accounted for in terms of a dichotomization of the internal and external environments. Minimally, two additional features need to be considered. The first is the social environment, which is a refinement of the external environment. The second is the genetic environment, which is a partial determinant of the internal environment.

The social environment is incredibly complex, as it encompasses all social interaction (or isolation) that an individual experiences throughout their life time, from birth to the instant before death. In addition, it encompasses the history of all previous social interactions that an individual has experienced. A full account of the social environment is far beyond the scope of this brief chapter. Such an account would include works from psychology, sociology, anthropology, communication sciences, and economics.

To better understand how the social environment can exert an influence on behavior, consider the following example. Suppose an individual is hiking in a remote wilderness area by himself and feels a need to urinate. It is quite likely that he will simply step off the trail and void his bladder. Now consider the same individual walking down Fifth Avenue in New York City. When he feels the exact same bodily sensation, he will act in a very different fashion. Instead of seeking immediate relief, he will seek out a public restroom. The main difference in the situation is the social environment. In the first case he is isolated from all others and in the second he is surrounded by a sea of humanity.

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00025-5

Consider the myriad genetic influences on our internal environments. Imagine twins. One of the twins has a genetic makeup that has rendered him or her unable to experience the taste we commonly refer to as bitter. The other twin does not have this genetic anomaly and can fully appreciate the taste of bitter. Now suppose both are equally thirsty (internal environment), but the only water available is laced with quinine, which makes it very bitter. It would likely be unpalatable to the twin with the ability to normally experience bitter but the twin with altered genetic makeup will have no trouble slaking his or her thirst with the quinine-laced water.

All four of these arenas – genetic, internal, social, and external environments – constantly interact to determine an individual's behavior on a moment-by-moment basis over the course of their life. These influences are depicted in Fig. 25.1. This results in an incredibly complex system of behavioral determinants in which an individual's genetic predispositions, current physiological and affective states, the current external environment, and the social context – both current and historical – interact to determine an individual's current behavior. Consider also that many of these factors can change over time, some very rapidly, and you begin to appreciate how daunting a task it is to try and account for even the simplest of behaviors. Now imagine trying to understand a behavior like drug use, as opposed to a simple behavior like putting on a coat when it is cold outside, and the complexities seem almost insurmountable.

The challenges notwithstanding, providing such an account is critical if we are to understand, prevent, and treat addiction. Efforts to understand this complex problem are largely the realm of the experimental and clinical behavioral pharmacologist. This multidisciplinary field draws on the expertise of many disciplines including genetics, epigenetics, molecular biology, biology, psychology, sociology, anthropology, and economics, among others. In addition, bidirectional approaches to understanding the determinants of drug abuse, in which those afflicted with addiction and their friends and families participate, are key to developing the fullest understanding of the determinants of this behavior as possible.

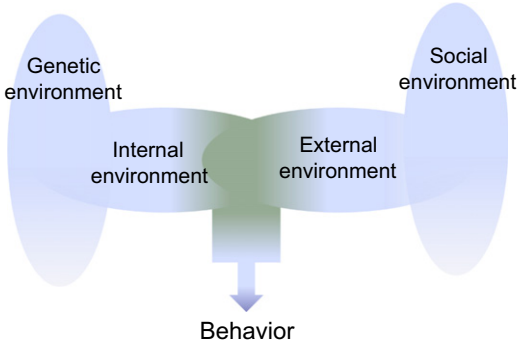


FIGURE 25.1 This schematic represents the continuous interplay between genetic makeup, social context, internal and external environments, and the production of behavior. As indicated in the figure, all four environments can influence each other.

A useful conception for gathering all of the various influences depicted in Fig. 25.1 into one rubric is a reinforcement and punishment categorization. This idea is not new. The roots of the idea can be traced at least as far back as Bain and the Scottish (and British) philosophers who noted how behavior tended to congregate around those things that satisfied a biological need (e.g. a presumably hungry lamb tended to organize its behavior around nursing). The relationship between behavior and its consequences was further described by Thorndike who enshrined it in his Law of Effect, which roughly states that those behaviors that produce satisfaction tend to increase in frequency while those behaviors that result in frustration tend to decrease in frequency. For example, Bain's hungry lamb learns how to nurse from his mother's udder because it results in the delivery of food, whereas nursing on her tail does not. Skinner, in one of the most elegant and productive careers in the field of psychology, refined this work into what is the modern field of the experimental analysis of behavior. Central to this work is the definition of reinforcement and punishment.

A reinforcer is a stimulus that, when presented after the occurrence of a behavior, increases the likelihood of the behavior being repeated in a similar context. A punisher is the opposite: a stimulus that, when it follows a behavior, results in a decreased likelihood that the behavior will be emitted in a similar context in the future. For example, if you have a garden gate and its latch does not work well so that you have to shake the gate to release the latch and permit you to open the gate (the reinforcer), you are likely to try the same shaking behavior next time you want to gain access to your garden. However, if shaking the gate results in you getting a splinter (punisher), you are less likely to try the same behavior in the future. For much more detailed descriptions of reinforcement and punishment, please consult the Further Reading list. The remainder of this chapter focuses on reinforcement for simplicity's sake; however, many of the same principles apply to punishers.

Reinforcement (and punishment) can be parsed further into both negative and positive varieties. In the garden gate example above, the opening of the gate is an example of a negative reinforcer because it results in the removal of a barrier (the gate) between you and something you want (access to the garden). If once you are in the garden you pick a strawberry and eat it because you are hungry, the strawberry is a positive reinforcer.

Reinforcement (and punishment) can be augmented (or degraded) depending on the contextual associations that surround it. For example, imagine you are fishing and having no luck. You are hot and thirsty. You decide to open a cold soda of a variety you have never tasted before and drink it. As soon as you begin to drink the soda, you have a strike and catch a fish. Now suppose this pairing of events (drinking and catching a fish) happens several times. An association will form such that the reinforcing efficacy of the soda is greater than could be accounted for by its thirst-quenching properties alone. Instead, it has acquired additional reinforcing efficacy through its repeated association with the catching of fish. In the parlance of behavior analysis, the soda is said to now possess primary reinforcing qualities (thirst quenching) and conditioned reinforcing efficacy (fish catching associations). My group has studied this process in human laboratory settings and further reading on the topic is included in the Further Reading list.

While it is not difficult to understand the environmental factors that pressure us to open a garden gate and eat a strawberry or that compel us to put on a coat when it is cold, it is difficult to understand why humans do something as complex as abuse drugs.

The observation that drugs of abuse could function as positive primary reinforcers opened the door to applying a behavioral account of behavior to addiction. Most drugs of abuse activate the same neural circuitry as nondrug primary reinforcers. In our earlier example, eating a strawberry when hungry served as a primary reinforcer. This behavior activated certain brain pathways, often referred to as the reward pathway. This pathway is associated with the receipt of a primary reinforcer (e.g. sexual activity, food, and water).

When a cocaine- or methamphetamine-dependent individual ingests a dose of either drug, the same neural circuitry is activated. The neural circuitry may become hyperactivated. This led Dr Leshner, former director of the National Institute of Drug Addiction, to term the process as a hijacking of a user's motivational system. The end result of such a process is that the individual begins to engage in compulsive drug use. This leads to a pattern of behavior in which drug seeking pushes everything else into the background. Many if not most people who experiment with a drug like cocaine or methamphetamine never develop a full blown addiction to the drug. This suggests that the process is somewhat malleable.

Drugs can also act as negative reinforcers by removing a noxious state. For example, consider a harried stockbroker. She often works 12-h days; many individuals rely on her to manage their retirement accounts on which their future depends. In addition, she has a family that is dependent on her for all the interactions associated with being a spouse and a mother. Now imagine that she makes a mistake one day and her clients lose the majority of their savings. Her family still expects her to be there for them, but her clients want her locked up because of her negligence. Her career is certainly over, at least as a stockbroker. Our unfortunate stockbroker could hardly be blamed for drinking excessively. But consider what the drinking does. It removes, at least temporarily, the stress of her life. In an intoxicated state, she may be able find respite from the anxiety and despondency she feels. In this sense, the alcohol can be considered to function as a negative reinforcer.

The same thing applies to the homeless man living on the streets of a large urban area. He has no place to sleep, no food, no clean clothes, and is probably sick. Moreover, there is a very high likelihood that he is physically assaulted on a routine basis. However, when he shoots up with heroin, it undoubtedly feels good. Like the stockbroker, he gains a respite from the despair he faces in his day-to-day existence. The heroin acts as a negative reinforcer.

Drugs may also function as conditioned reinforcers. Imagine a family with a 15-year-old daughter who has just moved into a new neighborhood. The young woman has left all of her friends behind and knows nobody in her new neighborhood. However, she desperately wants new friends. She meets a group of folks her own age; however, they all smoke marijuana when they are together. They are very uncomfortable with strangers around when they are doing this. They will not tolerate the new girl's presence unless she too smokes marijuana. So, in order to gain access to the reinforcer – the social group – she begins to smoke marijuana. In this case, the marijuana not only has the primary reinforcing efficacy attributable to its interaction with the young woman's reward pathway, it has also acquired properties of a conditioned reinforcer. It can be argued that this is one way in which addictive behavior is likely initiated.

The obvious complexity of obtaining a full understanding of the intricate environmental determinants of drug use cries out for an experimental line of work in which laboratory-based animal and human work influences human clinical trials. In the following sections each of these sources of data is considered. Entire books can and have been written on each of these topics, so only a flavor of the work is provided here. Selected additional readings are included in the Further Reading list.

LABORATORY STUDIES

A robust finding that has been observed by numerous investigators is that the density and magnitude of reinforcement in an environment is an important determinant of behavior. Work by Dr Nader and colleagues

(and many others) has demonstrated that when an animal (e.g. a primate) is given the opportunity to choose between self-administering cocaine and food, the number of choices made for cocaine varies as a function of the magnitude of the concurrently available food reinforcer. In this paradigm, the food is considered to be an alternative source of reinforcement to the cocaine. In a prototypic experiment, a monkey would be allowed to make repeated choices between a constant dose of cocaine and one pellet of food during an experimental session. In this case, the monkey might elect to take 80% of the available cocaine doses. In the next session, the monkey makes 10 choices between the same dose of cocaine and four pellets of food. In this situation, the monkey chooses cocaine on only 50% of the trials. Finally, a session occurs in which the monkey chooses between the same dose of cocaine and eight pellets of food. In this arrangement, the monkey elects to administer cocaine on only 20% of the trials. This preparation and series of experimental conditions demonstrates, in an elegant fashion, how the availability of an alternative source of reinforcement – when available in a sufficient magnitude – can moderate the reinforcing efficacy of a drug. This is shown by the reduced proclivity to self-administer the drug in the presence of increasingly large magnitudes of alternative reinforcement. This pattern of results has been obtained by many investigators and for many different drugs of abuse and animal species. Examples of this line of work are included in the Further Reading list.

Dr Stephen T. Higgins and colleagues were among the first to demonstrate the above principle in a human laboratory preparation. In a human laboratory procedure closely based on prior animal work, Dr Higgins and colleagues demonstrated that the same principle – high-magnitude alternative reinforcement competes with drug self-administration – pertained to humans. They recruited recreational cocaine users to take part in an outpatient study. Participants were allowed to make 10 repeated choices between 10 mg unit doses of intranasal cocaine and money. In this preparation, money was the alternative reinforcer. When the monetary magnitude was low (i.e. $0.05), participants selected cocaine almost exclusively, when the amount of available money per choice increased to $1.00 about 50% of the choices were for cocaine. When the amount of available money was further increased to $2.00 the choices were almost exclusively for money. As in the animal study, drug self-administration decreased as the magnitude of the available alternative source of reinforcement increased. In other words, cocaine users were able to readily forego cocaine self-administration when they had to choose between cocaine and a sufficiently high magnitude of alternative reinforcement. Many investigators have essentially replicated this work. In one project, we used methamphetamine users who met the diagnostic criteria for methamphetamine use disorder (not recreational users) and the same relationship held. High-magnitude alternative sources of reinforcement decreased the reinforcing efficacy of the methamphetamine.

The above examples demonstrate one way in which behavior can be altered by engineering specific environments. Essentially, the reinforcer-directed behavior of drug self-administration was altered by utilizing external environmental manipulations that decreased the reinforcing efficacy of the drug. There are other ways to influence a drug's reinforcing efficacy and the self-administration of the drug. As Fig. 25.1 illustrates, any of the four environments (genetic, internal, external, and social) could potentially alter the drug-taking behavior of an individual. In the above examples, the behavior was altered by manipulating the external environment directly (i.e. arranging forced choices between the drug and different magnitudes of salient alternative sources of reinforcement). Other commonly investigated methods include the use of pharmacotherapeutic approaches that alter the internal environment. For example, nicotine replacement therapy occupies the internal nicotinic receptors and decreases the compulsion for the individual to self-administer nicotine. Another promising line of work is to utilize an immunotherapeutic approach to interfere with drug's ability to exert its effect. Many psychosocial procedures, such as relapse prevention, seek, in part, to alter the drug user's social environment so that they remove themselves from situations that have historically been associated with drug use.

All of these procedures have a goal of reducing a drug's reinforcing efficacy in common. This in turn reduces the likelihood that the drug will be consumed. This reduction of reinforcing efficacy is the hallmark of successful drug abuse treatment paradigms. Arguably, the most successful treatment paradigm is contingency management, a procedure that provides salient reinforcement to treatment-seeking individuals contingent on drug abstinence. This is a procedure that, when successful, alters the drug user's external and social environments in such a fashion as to decrease the drug's reinforcing efficacy.

CONTINGENCY MANAGEMENT

One could make a strong case that contingency management has the strongest basic science foundation of any behaviorally based substance abuse treatment. In brief, contingency management for the treatment of substance use disorders is a procedure that decreases the reinforcing efficacy of a drug via the delivery of

reinforcement contingent on abstinence and/or the delivery of punishment contingent on drug use. Given that it is widely accepted, and has been recognized for decades, that drugs of abuse function as potent positive reinforcers, a procedure designed specifically to decrease the drug's reinforcing efficacy, and hence the control the drug will exert over an individual's behavior, has much to recommend it from a theoretical framework.

Contingency management interventions have been developed to decrease substance use and improve treatment attendance through the judicious application of behavioral principles. Contingency management is based on a robust basic science literature supporting a position that drug use is, in part, a form of operant behavior. The availability of alternative nondrug reinforcers (i.e. vouchers, prizes) should decrease drug use if they are available in sufficient magnitude and according to a schedule that is incompatible with substance use. Contingency management is a powerful technique that has been used effectively to promote abstinence from alcohol, benzodiazepines, cocaine, nicotine, opiates, marijuana, and methamphetamine. Three meta-analyses support contingency management efficacy. A meta-analysis of controlled studies of psychosocial treatments for addiction (published in the *American Journal of Psychiatry*) observed that contingency management interventions demonstrate the largest reductions in drug use when compared with other treatment modalities. In a separate meta-analysis, contingency management treatment resulted in successful treatment episodes 61% of the time, while comparison treatments resulted in successful treatment episodes 39% of the time. Contingency management has demonstrated reductions in drug use that persist for 12–18 months after completion of the contingency management intervention in some randomized clinical trials. The long-term effects of contingency management on drug use have been found to be comparable with the long-term effects of cognitive behavioral therapies in the treatment of cocaine dependence. Research is currently underway that is focused on determining schedules and/or durations of reinforcement that maximize the long-term effect of contingency management interventions on drug abstinence.

The efficacy of contingency management has been demonstrated in community clinic settings in large randomized clinical trials. Contingency management was included as a recommended treatment in guidelines published by the National Institute for Health and Clinical Excellence in the United Kingdom. Investigations of contingency management dissemination are currently underway, including studies designed to better understand systemic and clinical variables that impede and facilitate contingency management implementation. Novel approaches have been developed to assist in funding contingency management interventions, such as providing greater control of disability benefits and providing opportunities for employment contingent on drug abstinence.

In a prototypic contingency management intervention, drug abusers come to a clinical setting and provide a urine test, which is analyzed for drug metabolites three times per week. Every time the urine test is negative (e.g. no recent drug use has occurred), participants receive a voucher worth a certain monetary value. The value of these vouchers escalates for consecutive instances of abstinence and resets to an initial value for failure to abstain. Treatment typically lasts 3 months but recent work from our group has shown that longer duration treatments are associated with greater rates of posttreatment abstinence. Vouchers can be exchanged for goods or services that are in line with the development of a drug-free lifestyle. For example, a mother in treatment might use her vouchers to buy a toy for her young son. Another person may use vouchers to buy clothes for a job interview.

These procedures produce high rates of in-treatment abstinence and statistically and clinically significant rates of posttreatment abstinence. Positive outcomes are associated with higher magnitudes of reinforcement, minimal delays between providing urine tests and receipt of reinforcement, baseline severity of drug use (less severity at baseline predicts better treatment outcome), and longer duration of the intervention. Contingency management interventions are among the best studied means to clinically alter a person's environment in an attempt to reduce a drug's reinforcing efficacy.

CONCLUSION

Behavior is determined by multiple influences. This brief chapter presents a framework that expands conventional accounts of environment to include those factors internal to an individual. It is my belief that such an expansion is necessary if we are to fully appreciate the complexities of human drug-taking behavior. I have also argued that multiple environmental interactions influence a person's moment-by-moment behavior through a reinforcement (and punishment) process. With this as a starting point, it follows that those procedures designed to alter a drug's reinforcing efficacy will have the greatest treatment impact. Contingency management is arguably one of the most successful interventions for treating substance use disorders (especially psychostimulant addiction) and it clearly operates by rearranging the drug user's environment to alter the reinforcing efficacy of continued drug use. Many

additional facets such as epigenetic modulation of the internal environment, research into individual differences, and prevention studies are not considered in this chapter. A full account of human drug taking will also need to consider these topics.

SEE ALSO

Cocaine Addiction, Methamphetamine Addiction, Tobacco, Behavioral Economic Factors in Addictive Processes

Glossary

Punishment procedure an act is followed by a negatively valued stimulus or event.

Reinforcement procedure an act is followed by a positively valued stimulus or event.

Further Reading

Alessi, S.M., Roll, J.M., Reilly, M.P., Johanson, C.-E., 2002. Commentary on conditioned reinforcement. Experimental and Clinical Psychopharmacology 10, 101–103.

Carroll, M.E., Lac, S.T., Nygaard, S.L., 1989. A concurrently available nondrug reinforcer prevents the acquisition or decreases the maintenance of cocaine-reinforced behavior. Psychopharmacology 97, 23–29.

Dallery, J., Silverman, K., Chutuape, M.A., Bigelow, G.E., Stitzer, M.L., 2001. Voucher-based reinforcement of opiate plus cocaine abstinence in treatment-resistant methadone patients: effects of reinforcer magnitude. Experimental Clinical Psychopharmacology 9, 317–325.

Dutra, L., Stathopoulou, G., Basden, S.L., et al., 2008. A meta-analytic review of psychosocial interventions for substance use disorders. American Journal of Psychiatry 165, 179–187.

Higgins, S.T., 1997. The influence of alternative reinforcers on cocaine use and abuse: a brief review. Pharmacology and Biochemistry of Behavior 57, 419–427.

Higgins, S.T., Alessi, S.M., Dantona, R.L., 2002. Voucher-based incentives. A substance abuse treatment innovation. Addictive Behavior 27, 887–910.

Higgins, S.T., Badger, G.J., Budney, A.J., 2000. Initial abstinence and success in achieving longer term cocaine abstinence. Experimental and Clinical Psychopharmacology 8, 377–386.

Higgins, S.T., Bickel, W.K., Hughes, J.R., 1994. Influence of an alternative reinforcer on human cocaine self-administration. Life Science 55, 179–187.

Higgins, S.T., Budney, A.J., Bickel, W.K., 1994. Applying behavioral concepts and principles to the treatment of cocaine dependence. Drug and Alcohol Dependence 34, 87–97.

Higgins, S.T., Budney, A.J., Bickel, W.K., Foerg, F.E., Donham, R., Badger, G.J., 1994. Incentives improve outcome in outpatient behavioral treatment of cocaine dependence. Archives of General Psychiatry 51, 568–576.

Higgins, S.T., Budney, A.J., Bickel, W.K., Hughes, J.R., Foerg, F., Badger, G., 1993. Achieving cocaine abstinence with a behavioral approach. American Journal of Psychiatry 150, 763–769.

Higgins, S.T., Delaney, D.D., Budney, A.J., Bickel, W.K., Hughes, J.R., Foerg, F., Fenwick, J.W., 1991. A behavioral approach to achieving initial cocaine abstinence. American Journal of Psychiatry 148, 1218–1224.

Higgins, S.T., Silverman, K., 1999. Motivating Behavior Change among Illicit-Drug Abusers: Research on Contingency Management Interventions. American Psychological Association, Washington, DC.

Lussier, J.P., Heil, S.H., Mongeon, J.A., Badger, G.J., Higgins, S.T., 2006. A meta analysis of voucher-based reinforcement therapy for substance use disorders. Addiction 101, 192–203.

Nader, M.A., Woolverton, W.L., 1991. Effects of increasing the magnitude of an alternative reinforcer on drug choice on a discrete-trials choice procedure. Psychopharmacology 105, 169–174.

Petry, N.M., 2000. A comprehensive guide to the application of contingency management procedures in clinical settings. Drug and Alcohol Dependence 58, 9–25.

Prendergast, M., Podus, D., Finney, J., Greenwell, L., Roll, J., 2006. Contingency management for treatment of substance use disorders: a meta-analysis. Addiction 101, 1546–1560.

Rachlin, H., 1991. Introduction to Modern Behaviorism, third ed. W.H. Freeman and Company, New York, NY.

Roll, J.M., Higgins, S.T., 2000. A within-subject comparison of three different schedules of reinforcement of drug abstinence using cigarette smoking as an exemplar. Drug and Alcohol Dependence 58, 103–109.

Roll, J.M., Higgins, S.T., Badger, G.J., 1996. An experimental comparison of three different schedules of reinforcement of drug abstinence using cigarette smoking as an exemplar. Journal of Applied Behavior Analysis 29, 495–504.

Thompson, T., Schuster, C.R., 1968. Behavioral Pharmacology. Prentice Hall, Englewood Cliffs, NJ.

CHAPTER

26

Behavioral Economic Factors in Addictive Processes

James G. Murphy[*], *Christopher J. Correia*[$], *Ashley A. Dennhardt*[*]

[*]University of Memphis, Memphis, TN, USA, [$]Auburn University, Auburn, AL, USA

OUTLINE

Most modern theories of addiction view drug self-administration as an operant behavior that is maintained by the reinforcing properties of drugs. Widely abused drugs such as nicotine, heroin, and ethanol reliably reinforce the behavior that leads to their administration. A key task of any theory of drug use, however, is to predict the conditions in which drugs will be highly preferred or valued reinforcers. Behavioral economic theory uses the term reinforcing efficacy (RE), which in laboratory settings is quantified by the amount of behavior (e.g. lever presses, time) allocated to gain access to the reinforcer, to describe the relative level of preference for a reinforcer such as alcohol or drugs. According to behavioral economics, the reinforcing value of a given drug is a dynamic and contextually determined product of the direct reinforcing effects of the drug, individual difference factors related to decision-making (e.g. impulsivity), and the availability of alternative reinforcers, rather than a fixed property of the drug or an immutable characteristic of the individual. The following sections describe several key features of behavioral economic theories of addiction. After this general overview of behavioral economic models of addiction, the article ends with a review of studies that have specifically examined the role of several key behavioral economic variables in the development, progression, and cessation of substance use: drug price, the RE of drugs relative to alternatives, and manipulations of substance-free reinforcement.

EFFECT OF AVAILABILITY AND PRICE OF DRUGS AND ALTERNATIVE REINFORCERS

Behavioral economic theory predicts that the primary environmental influences on drug use are both constraints on access to drugs and the availability and value of alternative substance-free sources of reinforcement. Drug use is most likely when there are minimal constraints on drugs and substantial constraints on

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00026-7

access to valued nondrug reinforcers. Indeed, both naturalistic and laboratory research indicate that rates of substance use are highly sensitive to changes in the response cost or price associated with drug use; high rates of substance use are most likely in contexts devoid of substance-free sources of reinforcement, and substance use will generally decrease if access to alternative reinforcers is increased. It is important to note, however, that although several studies have identified alternative reinforcers that substitute for drug use (an inverse relation between levels of drug use and levels of the drug-free activity), drugs and alternative reinforcers can also be independents (no relation between levels of drug use and levels of the drug-free activity) or complements (positive relation between levels of drug use and levels of the drug-free activity) and this relation may be moderated by other variables. For example, whereas social activity has been shown to be a substitute for substance use in many adult samples, among young adults levels of socializing are often positively correlated with drinking, indicating a complementary relation.

It is important to point out that behavioral economics is a molar theory, which means that its goal is to account for patterns of behavioral allocation toward drugs and alternatives over time, rather than discrete instances of drug use. Drug use is assumed to vary predictably in relation to the relative availability of drugs and alternatives, but no predictions are made about whether an individual will use drugs in a given instance. In contrast, biological, social learning, and cognitive theories of addiction focus on proximal explanations for specific instances of drug use (e.g. craving, expectancies, peer pressure).

Changes in Relative Value as a Function of Engagement in These Activities

A key assumption of behavioral economic theories of addiction is that the cost–benefit ratio of drug consumption versus engagement in other activities changes as a function of the amount of engagement in these respective activities. The nature of the change in the cost–benefit ratio as a function of consumption (or participation), however, is markedly different for drugs versus many potential substitute activities. Drugs are viewed as price-habituated activities, which means that there is a negative relation between the amount of consumption over time and the benefit derived from a given consumption episode. This is consistent with the observation that tolerance reduces the rewarding effects of a given drug dose in experienced drug users. In contrast, many substance-free activities are price-sensitized, which means that there is a positive relation between participation and the benefits derived from the activity. Interpersonal relationships, exercise/sports, and hobbies such as knitting and playing bridge are examples of activities that will generally result in increased benefits over time, as intimacy (in the case of relationships) and skill generally increases with repeated engagement in these activities. If these activities are neglected, conversely, their cost–benefit ratio will increase (e.g. running after a period of immobility, attempting to access social support after a period of social isolation, attending statistics class after missing the previous 3 classes).

Behavioral economic theories suggest that the central dynamic of addiction is the relative distribution of behavior over time to constructive price-sensitized activities versus drug use. Because drug use is price-habituated, it is in theory a self-limiting activity. Whereas moderate substance use (e.g. a glass of wine with dinner, or several drinks during a weekly happy hour) results in reliable benefits (euphoria, relaxation, social facilitation), regular heavy drug use generally will diminish the value of any single unit of use, perhaps due to physiological or behavioral tolerance and perhaps mounting negative consequences that begin to tip the cost–benefit ratio. Indeed, the fact that the number of moderate drinkers and drug users far exceeds the number of heavy users and that even many substance abusers have periods of moderate use is consistent with this self-limiting prediction.

Behavioral economic theory predicts that there are several conditions that could result in increased drug use. First, an event that results in a reduction in substance-free substitutes, such as the loss of friend, family member, or job, moving to a new environment, or an injury or illness. Second, an event or situation that results in a reduction in the monetary or behavioral price of drug use. For example, for many young adults the college environment drastically reduces the total price of both alcohol and drug use, which includes more than just monetary cost. Alcohol and drugs are readily available at bars and parties, reducing the cost of time and effort to procure them, and the parental sanctions that served to increase the price of drinking or drug use during high school are generally absent. Moreover, many students have ample free time and are able to avoid taking morning classes, factors that minimize the extent to which heavy drinking results in functional impairment. All of these environmental factors may partially explain the fact that although as many as 40–50% of the US college students report frequent heavy drinking, most of these students substantially reduce their drinking in the years after graduation. Similarly, soldiers in Vietnam, an environment in which drugs such as heroin and marijuana were readily available without the threat of legal sanctions, reported high rates

of drug use and dependence but most stopped or reduced their consumption after returning to the United States, despite the fact that few participated in formal drug treatment. Other examples of situations that result in a reduction in the price of drinking and or drug use include turning 21 and being able to purchase alcohol legally, beginning a relationship with a heavy drinker or drug user, and moving into a neighborhood with a high density of drug dealers and liquor stores.

Because drug use and many substance-free activities are mutually substitutable, an increase in drug use related to these price reductions will often result in decreased engagement in substance-free activities. This decrease will in turn reduce the value of these price-sensitized, substance-free activities, resulting in greater behavioral allocation toward substance use. Thus, although the value of any individual occasion of substance use will continue to decrease as use escalates, the concurrent reduction in the value of price-sensitized, substance-free substitutes will often make drug use the option with the greatest immediate benefit. This is especially true for drugs of abuse that produce strong withdrawal effects (e.g. sedatives and opiates), which result in negatively reinforced patterns of use. This progression toward greater drug use is exacerbated by the fact that drug use often has a direct negative impact on other activities (e.g. hangover or withdrawal-related impairment, substance-related arguments, social stigma, legal fines/sanctions) in addition to the price-sensitized effect associated with diminished engagement in substance-free activities. Moreover, recent neurobiological research indicates that many drugs of abuse have direct neurochemical effects that exacerbate this narrowing of the behavioral repertoire toward drug use. Acute drug use has been shown to reduce the dopamine reward threshold such that dopamine firing occurs in response to milder rewarding stimuli. This might lead to greater pleasure or RE associated with activities that occur during a drug or alcohol use episode (e.g. creative activities, socializing, sexual activity). Chronic drug use, conversely, raises the dopamine reward threshold, making substance-free activities less reinforcing and less likely to effectively substitute for drug use.

It is also possible, however, for contextual events – such as an increase in the price of drugs or a decrease in the price of substance-free activities – to decrease the cost–benefit ratio of substance-free activities, leading to reduced drug use. In the examples cited above, graduating college and assuming adult responsibilities such as full-time employment and or beginning a family, or returning to the United States after military deployment, might be events that generate such a change in the cost–benefit ratio. The efficacy of contingency management interventions – a substance abuse treatment that provides positive reinforcement for verified periods of drug abstinence – suggests that even highly dependent substance abusers are sensitive to changes in the cost–benefit ratio of drug and alternative reinforcers. Furthermore, research has demonstrated that individuals who successfully recover from substance abuse without participating in treatment often note increases in rewards from key life domains (work, family, hobbies), which suggests that substance abuse patterns are sensitive to contextual events that change the relative availability of drug-free rewards.

Important Temporal Influences on Choice between Drugs and Alternatives

The choice dynamic described above is critically influenced by the temporal context of the decisions. Drug choices will often maximize utility (i.e. happiness, satisfaction) if the time frame for reward maximization is short. Drinking heavily at a bar on one particular night might maximize utility for that evening compared with staying home and studying. If the temporal context is extended (e.g. 1 month, or year), however, behavioral allocation toward constructive, price-sensitized activities such as studying would maximize utility. Behavioral economics is a molar theory of choice and as such assumes that addiction entails a series of "distributed choices" (e.g. drinking vs abstaining on a given night) rather than a discrete choice (to be a heavy drinker vs an abstainer) and that the frame of reference used to estimate the relative value of a series of choices concerning drug use will determine the amount of drug use over time.

Available data suggest that there are interspecies, developmental, and individual differences in the extent to which decisions are made on the basis of relatively short-term versus long-term outcomes. For example, although the value of all rewards decreases as their receipt is delayed (i.e. delay discounting), there are substantial individual differences in the degree of discounting of delayed rewards, and behavioral economic theory suggests that this discounting phenomenon may be a core feature of addiction. That is, consistently using drugs instead of engaging in activities with greater long-term benefit (e.g. work, relationships, school) may be due to sharp devaluing of these delayed rewards relative to the immediate benefits of drug use. Research with humans and animals suggests that relative preference for smaller, sooner rewards increases sharply as the receipt of the smaller reward is imminent. This choice pattern is well described by a hyperbolic discount function and is quite consistent with the choice dynamics associated with substance abuse, including switching preference from larger delayed rewards (e.g. academic or vocational pursuits, physical health) to substance use when drugs are immediately available.

Studies have found that a variety of substance-abusing groups (alcohol, cocaine, heroin, nicotine abusers, pathological gamblers) discount delayed monetary rewards more steeply than matched controls. Research examining delay discounting in both human and nonhuman animals suggests that delay discounting can be exacerbated by the effects of chronic drug use and drug withdrawal, and might even improve after extended abstinence. Delay discounting also seems to be an etiological risk factor that contributes to the development of substance abuse and has implications for the likelihood of change after treatment. One compelling animal study found that rats that showed a strong preference for smaller, immediate amounts of food compared with larger, delayed food rewards subsequently showed greater levels of cocaine use. Among humans, discounting may predict the acquisition of tobacco use, and nonsubstance-abusing teens with a family history of substance abuse report higher discounting than similar teens without a family history of substance abuse. Studies with smokers and alcohol-dependent individuals indicate that steep discounting at baseline predicts relapse or poor treatment response. Finally, although delay discounting rates are fairly stable over time, the state versus trait nature of discounting remains unresolved, and researchers have begun to identify techniques to extend an individual's time horizon.

The Influence of Price on Drug Use

As predicted by the Law of Demand, consumption of most commodities, including alcohol and other psychoactive drugs, is inversely related to price or response requirement. However, there are cross-commodity and individual differences in the relative price sensitivity or elasticity of consumption. Demand for a reinforcer is unit-elastic when increases in price lead to proportional decreases in consumption (resulting in no change in overall expenditures on that reinforcer) and elastic when increases in price lead to greater than proportional decreases in consumption (resulting in a reduction in overall expenditures on that reinforcer). Demand for nonessential and/or luxury items, such as restaurant meals, is generally elastic. Alternatively, demand for a reinforcer is inelastic when increases in price lead to less than proportional decreases in consumption, resulting in an increase in overall expenditures on that reinforcer (e.g. demand for gasoline during the summer travel season). It is worth noting that the elasticity of a good may be partially moderated by the income of the consumer. In behavioral terms, income is defined as the total amount of money, time, energy, or other resources that can be allocated to various reinforcers. The behavior of consumers with a relatively high income may be less sensitive to price increases than those with relatively low income. Thus, individuals with large amounts of free time (e.g. college students, unemployed adults) are more likely to "spend" large amounts of time procuring and using drugs.

In considering the effects of price on consumption, it is important to note that the behavioral economic definition of price goes beyond the simple monetary cost. Chaloupka and Pacula provided a comprehensive review of a variety of factors that can influence decisions about smoking, including the monetary costs of purchasing cigarettes, restrictions on where and when people can smoke, the fines and other legal consequences of smoking at unauthorized locations, and awareness of the short- and long-term health risks associated with smoking. More generally, the full price of any good, service, or reinforcer consists of four basic components: (1) monetary cost, (2) time and effort costs, (3) potential legal cost, and (4) potential health cost. Increases in any of these costs can lead to decreases in substance use, and, conversely, lowering the cost can lead to increases.

Several laboratory studies have modeled the relationship between price and substance consumption. One study used a simulated alcohol purchase task to assess levels of alcohol consumption across a range of prices in a sample of social drinkers. The average number of standard drinks consumed was approximately seven when the price was $0.25 or less per drink, remained at or above five drinks at prices up to $1.50 per drink, and then became more elastic and showed a steady linear decrease as prices increased. Not surprisingly, participants with a history of engaging in heavy episodic drinking were willing to spend more to consume alcohol, demonstrating that individual difference factors such as history of use and dependence can play a role in the demand elasticity of alcohol. A similar purchase task was used to assess cigarette consumption in adolescents. The authors found that while hypothetical cigarette purchases were elastic, the majority of adolescents would purchase at least one cigarette at prices up to $2.50 and 20% will still purchase a cigarette at $7.00. Another study that used a sample of 200 heavy drinking college students revealed that students' hypothetical drinking decreased across the entire range of drink prices if they were instructed to imagine that they had a college class the next morning, and decreased further if they were instructed to imagine that the morning class included a test. The next morning class can be viewed as an alternative reward or as an indirect method of increasing the "cost" of drinking; the results provide support for the behavioral economic view that decisions to drink or use drugs are sensitive to price and alternative reinforcement contingencies.

An abundance of population-based research also supports the notion that an increase in the monetary price of alcohol leads to decreases in consumption,

making demand for alcohol fairly elastic. However, recent studies suggest that the relationship between price and consumption is quite complex. An analysis of Swedish alcohol price and sales data collected between 1984 and 1994 confirmed that increases in price led to decreases in consumption for spirits, wine, and beer, but also revealed that increases in price led some consumers to switch to lower-quality (and lower priced) products so that they could maintain their level of consumption.

The RE of Drugs Relative to Alternative Reinforcers

Behavioral economic theory views addiction as a state in which the RE of drugs is high compared with the RE of available drug-free alternative activities. In laboratory settings, the RE of drugs is generally measured by the levels of operant responding for a drug, the quantity of the drug reinforcer earned or consumed during the session, the extent to which drug-reinforced responding is sensitive to increasing price or schedule requirement, or the proportional response rate toward drugs. Herrnstein's matching law ($\log B_1/B_2 = a \log r_1/r_2 + \log c$), for example, states that the proportional resource allocation directed toward available activities (B_1/B_2) equals the proportion of reinforcement obtained from the activities (r_1/r_2). The a and c parameters reflect sensitivity to reinforcement frequency and bias for one or the other alternative, respectively. The matching law has accurately predicted choice in numerous laboratory studies involving a variety of species and reinforcers, including studies examining drug administration. RE measures that plot the sensitivity of drug consumption and drug-reinforced responding to increases in price (demand curves) have also been used to predict the abuse liability of drugs.

There are several measures of RE that can be administered to human participants in clinical settings, and several recent studies provide initial support for the validity of these novel measures of substance abuse problem severity. Tucker and colleagues developed a measure of RE based on proportional discretionary monetary expenditures to alcohol. Their discretionary expenditure measure, which is modeled after laboratory paradigms that consider response output and consumption as distinct facets of RE, is based on the premise that alcohol is a more valued reinforcer for an individual who allocates 45% of her expenditures to alcohol consumption than it would be for an individual who allocates only 10% of her expenditures to alcohol consumption, even if these individuals had similar alcohol consumption levels. In two studies of natural recovery from alcohol dependence, Tucker and colleagues assessed the proportion of discretionary income allocated to alcohol in alcohol-dependent individuals before a quit attempt. The results indicated that relative resource allocation to alcohol predicted drinking outcomes, whereas traditional measures of consumption and dependence did not. Participants who relapsed within the 2-year follow-up period allocated a greater proportion of their money to alcohol in the year before the attempted resolution, even though their levels of alcohol consumption and problems during this period were similar to resolved participants. Thus, consistent with behavioral economic predictions, RE measures based on resource allocation seem to have utility in predicting clinically relevant changes in drinking.

Another way to measure reinforcement is to assess the frequency of specific, rewarding activities. Reinforcement surveys such as the Pleasant Events Schedule and the Adolescent Reinforcement Survey Schedule measure the frequency of participation in and subjective enjoyment derived from various activities (e.g. eating good meals, meeting someone new, gardening). The cross product of the frequency and enjoyment ratings provides an approximation of obtained reinforcement (e.g. the sum of recent time spent in enjoyable activities). Participants make two sets of frequency and enjoyment ratings for each item – one for drug-free activities and one for activities that take place while using drugs or alcohol. The substance-related reinforcement score is a measure of drug-related enhancement of other activities (e.g. socializing while high) rather than a direct index of the reinforcing effects of drug use (e.g. sedation, euphoria). The substance-free cross product score provides an index of substance-free reinforcement. The primary score derived from the modified reinforcement surveys is the ratio of substance-related reinforcement to total reinforcement, an index of the RE of substance use. Correia and colleagues derived this index from Herrnstein's matching law: substance-related reinforcement/(substance-free reinforcement + substance-related reinforcement). Cross-sectional studies with college drinkers and psychiatric patients demonstrated that the ratio measure predicted unique variance in drinking quantity beyond the effect of substance-related reinforcement alone.

Another study examined whether this reinforcement survey measure of RE was related (prospectively) to drinking outcomes among heavy drinking college students who completed a brief alcohol intervention. Only women showed a significant reduction in drinking at the 6-month follow-up, and RE accounted for unique variance in their drinking outcomes. Women who derived a smaller proportion of their total reinforcement from substance use at baseline reported lower levels of follow-up drinking, even after controlling for their baseline drinking level. Thus, individuals

who have a number of enjoyable alternatives to drinking may have an easier time reducing their consumption after an intervention. Men and women who reduced their drinking by at least five drinks per week showed increased proportional reinforcement from substance-free activities at follow-up, which suggests that successful drinking reductions are associated with more global changes in lifestyle and activity participation.

Another analysis from this treatment trial examined a reinforcement measure derived from a hypothetical alcohol purchase task wherein participants report the number of drinks they would purchase across 14 different prices ($0–$13.00). Reported consumption and expenditures are plotted as a function of price (demand curves), which are in turn used to generate several indices of RE, such as maximum level of alcohol consumption at low cost (intensity of demand), maximum level of alcohol expenditure (O_{max}), and several measures that reflect the degree to which consumption decreases with increasing price (breakpoint, P_{max}, and elasticity of demand). Murphy and colleagues examined the psychometric properties of the indices derived from the alcohol purchase task and found them to correlate with alcohol consumption, alcohol-related problems, and other RE variables, demonstrating good construct validity, and to have good test–retest reliability at a 2-week interval. As hypothesized, a number of the facets of reinforcement predicted weekly alcohol consumption and heavy drinking at the 6-month postintervention follow-up assessment. Participants who at baseline reported greater maximum expenditure (i.e. O_{max}) for alcohol and lower price sensitivity (i.e. breakpoint, P_{max}, and elasticity) reported greater follow-up weekly drinking. Other studies have shown that greater intensity of demand is associated with more problematic use (i.e. alcohol use disorder symptoms, alcohol problems), which suggests that this index may be useful in identifying more severe patterns of use.

Little and Correia provided further evidence for the construct validity and potential utility of self-report reinforcement measures that utilize hypothetical alcohol purchases. Hypothetical alcohol purchases on the multiple choice procedure (MCP) were significantly correlated with measures of alcohol consumption and problems and with a laboratory paradigm that included real choices between alcohol and monetary amounts. Another recent study of adolescent smokers found that RE measures derived from hypothetical cigarette demand curves demonstrated significant associations with smoking and nicotine dependence. Thus, RE measures that are based on simulated drug purchases seem to be valid and to show meaningful relations to real-world patterns of substance use and problems. RE indices may provide unique information about the value or strength of preference for substance use that is not typically included in clinical or applied research contexts. Existing substance abuse measures provide important information on consumption levels and the immediate risks and harmful effects of drinking/drug use, but they do not describe the relative prominence of substance use in an individual's overall lifestyle or the amount of resources allocated to substance use, which may be more predictive of the course of substance use over time. Proportional reinforcement from substance use relative to substance-free activities operationalizes an important feature of substance misuse: devoting considerable time/resources to substance use and neglecting other important activities. This concept is included in the DSM-IV–dependence criteria but is not explicitly measured in traditional substance use assessment batteries. RE may be an especially useful means of discriminating between individuals with similar substance use patterns, but different levels of risk (and need for treatment) based on their overall pattern of behavioral allocation/reinforcement.

Drug Use Can Be Decreased by Increasing Drug-Free Reinforcement

The relationship between drug use and drug-free reinforcement has been studied through a combination of laboratory, naturalistic, and clinical studies. When studied in the laboratory, the dependent variable is typically a measure of drug use preference. Examples include the amount of the drug consumed, the amount of effort devoted to obtaining a drug, or the number of times the drug use is chosen over an alternative reinforcer. Independent variables involve manipulations of the availability, magnitude, and delay associated with an alternative reinforcer. For example, a series of studies conducted by Higgins and colleagues have shown that the reinforcing value of cocaine is malleable and highly influenced by the presence and magnitude of an alternative monetary reinforcer. Higgins, Bickel, and Hughes examined the influence of an alternative reinforcer on cocaine use among four adult humans. During 11 laboratory sessions, participants made repeated choices between cocaine versus placebo or between cocaine versus varying amounts of money ($0–$2.00). Participants always chose cocaine, rather than placebo, indicating that the drug served as a reinforcer. When subjects were asked to chose between cocaine and money, the amount of cocaine consumed decreased as the amount of concurrently available money increased. More recent studies by Correia and colleagues have used an MCP to investigate preference

for alcohol among college students. The results indicate that preferences are influenced by the immediacy of alternative monetary reinforcers. Thus, for both cocaine and alcohol, the preference for the drug decreases as the alternative monetary reinforcer is larger and more immediately available. The relationship has now been studied across a wide range of drugs (e.g. cigarettes, alcohol, cocaine, heroin), species (e.g. rats and mice, human and other primates), and human populations (e.g. drug-dependent users, nondependent users, college students). Across these laboratory studies using a variety of methodologies, participants generally show a greater preference for drug use when the value of the alternative reinforcer is small, the alternative reinforcer is delayed, or when the price of the alternative reinforcer is increased. These findings led Vuchinich and Tucker to conclude that drug use emerges as a highly preferred activity when constraints on the drug are minimal and alternative reinforcers are sparse or difficult to acquire.

The relationship between drug use and alternative reinforcers has also been studied and well documented in the natural environment. Several of these studies have used measures of behavioral allocation, such as the Pleasant Events Schedule and the Adolescent Reinforcement Schedule Survey, to measure engagement in a range of potentially reinforcing activities. A series of studies on college student drinking reported that the frequency, quantity, and negative consequences of alcohol use are predicted by the ratio of reinforcement derived from drug-related reinforcement to reinforcement derived from drug-related activities (e.g. school work, relationships, employment). These studies have also revealed that drug-free reinforcement is associated with increased motivation to change alcohol use; that increases in substance-free activities like exercise can lead to decreases in substance use; and that reduced drinking after a brief motivational intervention is associated with an increase in substance-free reinforcement. Studies using similar measures of behavioral allocation to study substance use among psychiatric outpatients and individuals enrolled in treatment for cocaine use disorders have produced consistent results. Interestingly, a recent study that examined substance-free reinforcement in heavy drinking college students found that frequent heavy drinkers reported greater substance-free reinforcement from peer and sexual activities. So, although overall substance-free reinforcement may increase after making a drinking reduction, several seemingly important reinforcement domains might suffer, at least among young adults, an outcome that might contribute to the overall low levels of treatment seeking and high rates of relapse. In general, the aforementioned studies suggest that alternative drug-free reinforcement is critically related to the development, progression, and cessation of substance abuse patterns.

BEHAVIORAL ECONOMIC INTERVENTION APPROACHES

The research reviewed above has influenced the development of several treatment approaches that explicitly attempt to increase the price of drug use and to increase access to alternative reinforcers. Contingency management (CM) is perhaps the most explicit example of a treatment designed to decrease substance use by increasing the density of drug-free reinforcement and the cost of substance use. As summarized by Higgins, all forms of contingency management, regardless of the substance(s) being targeted, the population, or the exact nature of the procedures, share a common conceptual framework, which calls for changes in the substance user's environment such that (1) drug use and abstinence are readily detected, (2) drug abstinence is readily reinforced, (3) drug use results in a loss of reinforcement, and (4) the density of reinforcement derived from nondrug sources is increased. As previously noted, contingency management procedures have been used effectively to reduce the use of a number of substances, including cigarettes, alcohol, cocaine, marijuana, and opiates.

The community reinforcement approach (CRA) is one example of a CM intervention. In the original Hunt and Azrin study, CRA participants were provided with counseling and material resources (e.g. telephones, televisions) to improve employment prospects, marital and family relationships, and nondrinking social and recreational interactions. These goods and services were provided to help clients access alternative sources of reinforcement by providing entertainment, or facilitating communication with employers and social partners.

Several other treatments highlight the various ways in which the reinforcing value of drug use can be decreased by increasing drug-free reinforcers. Several treatment approaches focus on the development of substance-free coping and social skills. These approaches are designed to address basic skill deficits that prevent clients from adequately coping with emotions and deriving satisfaction from a social-interpersonal situations at work, with friends, and with significant others. The coping skills training (CST) program, developed by Peter Monti and colleagues, focuses on delivering positive and negative feedback, listening skills, conversation skills, developing a sober support network, and assertiveness training. Behavioral activation therapy also promotes the development of skills to promote engagement in substance-free reinforcers. Consistent with the behavioral economic perspective the CST and behavioral activation serve to

decrease the reinforcing value of drug use by increasing the reinforcing value of drug-free social interactions and activities. Recent clinical outcome studies and meta-analysis continue to support the clinical efficacy of CM and other behavioral approaches. In addition to supporting the clinical utility of specific treatments, these treatment studies also provide additional evidence on the relationship between drug use and drug-free reinforcement.

CONCLUSIONS

Behavioral economic researchers have made considerable progress in describing the individual difference and contextual variables that influence the development, progression, and cessation of a variety of addictive behavior patterns. Key strengths of behavioral economic research include (1) a strong theoretical foundation in economic and learning theory, (2) a clear translation of concepts and measures from basic laboratory research on drug self-administration, including research with nonhuman animals to clinical settings, (3) a tradition of applying quantitative models to the description and prediction of substance abuse and associated decision-making processes, (4) decision-making tasks and paradigms that are amenable to both clinical translation (e.g. predicting treatment outcome) as well as to basic science translation (e.g. identifying genetic and fMRI correlates of discounting, studying interventions for discounting, or RE in animal models), and (5) concepts and experimental paradigms that have lead to the development of efficacious treatments such as CM. Key future directions include (1) the continued development of practical and valid measures of RE and delay discounting that could be used in clinical settings to predict need for and response to intervention, (2) the continued development of intervention approaches that are informed by behavioral economics, including brief and efficacious approaches for reducing delay discounting and increasing engagement in substance-free activities, and (3) efforts to understand the neural correlates of the destructive decision-making processes that underlie addiction. Regarding the final suggestion, the emerging field of "neuroeconomics" has the potential to contribute to the development of improved pharmacological and behavioral intervention that directly target and improve individuals' executive functioning and self-regulatory systems.

SEE ALSO

Impulsivity, Disinhibition, and Risk Taking in Addiction, Personality and Addiction Processes, Relapse and Lapse

Glossary

Behavioral economics a theory that integrates the principles and empirical methods from psychology and economics to understand how humans and animals allocate scare resources (behavior, time, money) under systems of constraint. Behavioral economics views the decision to use drugs as a cost–benefit analysis that is based on the relative availability, delay, and value of drugs versus alternative reinforcers.

Community reinforcement approach a behavioral approach to treating substance abuse problems that aims to reduce positive reinforcement for substance use and to increase positive reinforcement in social, recreational, familial, and vocational domains.

Complement a reinforcer acts as a complement when consumption of that reinforcer varies inversely with the price of another reinforcer. Coffee and sugar may function as complements such that an increase in the price of coffee may lead to a decrease in the consumption of sugar.

Contingency management a behavioral treatment approach that provides contingent positive reinforcement (vouchers, money) based on chemically verified abstinence for a given period of time (typically 1–3 days). Escalating schedules of reinforcement are often used to achieve prolonged stretches of abstinence.

Delay reward discounting refers to the level of decrease in the subjective value of a reinforcer as a function of the time until it is delivered; provides a behavioral economic index of impulsivity.

Demand curve a quantitative representation of consumption and response output (expenditures) for a given commodity (e.g. drug, food product) across a range of prices. Consumption and expenditures/response output are plotted on the *Y*-axis; price is plotted on the *X*-axis.

Intensity of demand maximum consumption level observed in a demand curve paradigm; consumption level at the lowest price increment.

O_{max} maximum expenditure or response output observed in a demand curve paradigm.

P_{max} the price associated with the maximum expenditure or response output (O_{max}) observed in a demand curve paradigm. P_{max} corresponds to the price at which demand becomes elastic.

Breakpoint the price at which consumption is completely suppressed in a demand curve paradigm.

Elasticity the rate of decrease in consumption as a function of price in a demand curve paradigm; demand is considered to be relatively elastic if a 1% change in price results in a greater than 1% change in consumption.

Price the total cost to acquire a good, service, or reinforcer which is the combination of (1) monetary costs, (2) time and effort costs, (3) potential legal cost, and (4) potential health and social costs.

Price-habituated activities when there is a negative relation between the amount of participation in the activity over time and the benefit derived from a given occasion of participation.

Price-sensitized activities when there is a positive relation between time spent engaging in the activity and the benefits derived from the activity.

Reinforcing efficacy the ability of a reinforcer to maintain high levels of operant responding. RE is generally measured by the levels of operant responding for a drug, the quantity of the drug consumed during a laboratory session, or the proportional response rate toward the reinforcer relative to alternative reinforcers.

Substitute a reinforcer that shares some essential property with another reinforcer, such that an increase in price in one reinforcer will lead to increased consumption of the substitute reinforcer (e.g. coffee and tea often function as economic substitutes).

Further Reading

Ainslie, G., 2001. Breakdown of Will. Cambridge University Press.

Bickel, W.K., Miller, M.L., Yi, R., Kowal, B.P., Lindquist, D.M., Pitcock, J.A., 2007. Behavioral and neuroeconomics of drug addiction: competing neural systems and temporal discounting processes. Drug and Alcohol Dependence 90 (Suppl.1), S85–91.

Camerer, C., 1999. Behavioral economics: reunifying psychology and economics. Proceedings of the National Academies of Sciences of the USA 96 (19), 10575–10577.

Carroll, M.E., Anker, J.J., Perry, J.L., 2009. Modeling risk factors for nicotine and other drug abuse in the preclinical laboratory. Drug Alcohol Dependence 104, 70–78.

Higgins, S.T., Heil, S.H., Plebani-Lussier, J., 2004. Clinical implications of reinforcement as a determinant of substance use disorders. Annual Review of Psychology 55, 431–461.

Loewenstein, G., 1999. A visceral account of addiction. In: Elster, J., Skog, O.J. (Eds.), Getting Hooked: Rationality and Addiction. Cambridge University Press, Cambridge, UK, pp. 188–213.

MacKillop, J., Miranda, R.M., Monti, P.M., Ray, L.A., Murphy, J.G., Rohsenow, D.J., McGeary, J.G., Swift, R.M., Tidey, J.W., Gwaltney, C.J., 2010. Alcohol demand, delayed reward discounting, and craving in relation to drinking and alcohol use disorders. Journal of Abnormal Psychology 119, 106–114.

Murphy, J.G., Correia, C.J., Vuchinich, R.E., 2009. The behavioral economics of substance abuse. In: Cohen, L.M., Collins, F.R., Young, A.M., McChargue, D.E., Leffingwell, T.R., Cook, K.L. (Eds.), Pharmacology and Treatment of Substance Abuse: Evidence and Outcome Based Approaches. Routledge, New York, pp. 505–528.

Rachlin, H., 1997. Four teleological theories of addiction. Psychonomic Bulletin and Review 4, 462–473.

Vuchinch, R.E., Heather, N. (Eds.), 2003. Choice, Behavioural Economics and Addiction. Pergamon/Elsevier Science, Amsterdam, The Netherlands.

CHAPTER

27

Cognitive Factors in Addictive Processes

*Clayton Neighbors**, *Kristen P. Lindgren*[$]*, *Lindsey Rodriguez**, *Judy Tidwell**, *Ivori Zvorsky*****

[*]University of Houston, Houston, TX, USA, [$]University of Washington, Seattle, WA, USA, [**]University of Richmond, Richmond, VA, USA

OUTLINE

INTRODUCTION

Attempting to define cognitive factors in addictive processes is a daunting task. One might argue that cognition is at the root of virtually every aspect of addiction, beginning with the decision to smoke a cigarette for the first time or to take a first sip of beer, to the successful navigation of occasional temptations encountered by a former substance abuser who has maintained abstinence for many years, and all of the points in between. There are many domains that might be considered in this context, including learning and memory, motives, intentions, attitudes, social perception, coping, self-awareness, self-control, self-regulation, attention, craving and urges, expectancies, attributions, cue-reactivity, neural networks, and implicit processes. We have elected to focus on a subset of these domains and are painfully aware that there are many other relevant topics which should and would be included in a more comprehensive treatment of the topic.

There are a number of contextual issues that warrant brief mention at the outset. First, it seems important to define what we mean by addictive processes. Second, it seems important to point out that processes may differ considerably depending on where an individual may be in the developmental course of addiction. Third, it is difficult to disentangle cognition entirely from emotion. Fourth, and finally, many of the cognitive processes underlying addiction (and behavior, more generally!) are not entirely conscious.

For the purposes of this chapter, we define addictive processes as the psychological and behavioral operations which relate to continued engagement in activities which an individual or society perceives as maladaptive, undesirable, and beyond the individual's control. Problematic substance use is a prototypical domain for discussing cognitive factors in addictions but similar processes may characterize a wide range of behaviors when taken to the extreme, including gambling, eating, shopping, internet use, sex, and so on.

Principles of Addiction, Volume 1
DOI: http://dx.doi.org/10.1016/B978-0-12-398336-7.00027-9

The developmental course of addiction may vary considerably across behaviors and individuals, but there are some common landmarks that seem universal. Ideally, all journeys have a beginning, a middle, and an end, and so it is with addiction. In the beginning, engagement in an addictive behavior is characterized by many positives and few negatives. For example, drinking is associated primarily with fun and social facilitation and gambling is often seen as a source of enjoyment and recreation. The middle of the journey may be divided into many subsections – characterized by increasingly disproportionate negative consequences and repeated attempts to disengage from the behavior. The end, in an ideal case, is a look in the rearview mirror, and a sigh of relief that one has survived a long and difficult journey. While it is the middle of the journey that typically receives the bulk of the attention in the research on cognitive factors and addictions, cognitive factors operate across the developmental course and will differ depending on an individual's place or stage in the journey. For example, cognitive processes related to self-control, cravings, and urges seem less relevant early and more evident later. In contrast, the opposite seems true for processes regarding perceptions of peer influence and positive expectancies. Furthermore, all along the way, cognition and emotion go hand in hand. A focus on cognitive processes may seem to suggest a focus on logical (or illogical) operations that are independent of affect. Human nature belies this possibility as many people's thoughts are about feelings and vice versa. Moreover, it is impossible to divorce how people think from how they feel. Thus, while our focus is on cognitive processes, it is important to keep in mind that affect is always intertwined with cognition. Finally, more and more work has begun to consider cognitive influences on addictive processes which operate outside of explicit awareness, often referred to as automatic or implicit processes. These are considered in some detail in the latter half of the chapter.

OUTCOME EXPECTANCIES

Among the broadest cognitive processes that have been shown to play an important role in addictive behaviors are those referred to as "outcome expectancies" or "expectancies." Most commonly studied with respect to alcohol, but more recently receiving increasing attention in studies of other drugs (e.g. cocaine, marijuana, tobacco) and addictive behaviors (e.g. gambling), expectancies refer to an individual's expectations of the outcomes associated with drug use. Theories of outcome expectancies reflect influences both from basic learning (e.g. positive and negative reinforcement) and cognitive theories. Expectancies are thought to reflect both an individual's past experiences of engaging in a behavior – that is, the extent to which one has experienced positive or negative reinforcement or punishment when using a drug – and that individual's expectations of the future consequences of engaging in that behavior. Thus, expectancies are argued to reflect both direct and indirect (vicarious) forms of learning that are, ultimately, stored as cognitive representations in memory. Whether expectancies stem largely from conscious or nonconscious cognitive processes is under debate, but, as discussed in more detail below, there appears to be general agreement that there is at least a significant nonconscious component to expectancies.

Expectancies are believed to develop from experience; thus, expectancies will vary as a function of the outcomes that an individual has experienced in conjunction with specific behaviors. Thus, a person who has smoked crack cocaine will likely have different expectancies about crack cocaine than an individual who has never tried it. Expectancies can also be derived from vicarious learning and observation of the results of behaviors performed by models (e.g. parents, peers, the media). For example, a child with limited observation of people consuming alcohol will have different expectancies than one growing up in a home where both parents drink heavily. Not only are there variations in expectancies across individuals but there are also variations within individuals. Thus, a child with limited exposure will have different expectancies as she ages and encounters more models and/or begins to have her own direct experiences with alcohol. Expectancies, both in terms of type (positive and negative) and intensity (weak vs. strong), have been shown to predict behaviors with stronger positive expectancies predicting greater use.

Expectancies have often been measured using self-report questionnaires with Likert-scale response options. Most scales include measures of both positive (e.g. social facilitation, tension reduction) and negative expectancies (e.g. increased aggression, cognitive impairment). Consistent with expectancy-value models, some scales also include an additional set of questions that ask respondents to rate the extent to which they view each outcome as positive or negative, which is designed to account for variance in how positively or negatively a given outcome is viewed by an individual. For example, some respondents may view an increase in aggression associated with drinking as a positive effect of drinking and/or a decrease in the ability to remember things as a positive outcome of marijuana use. Individual items within and across questionnaires also vary in the extent to which they: (a) focus on outcomes that affect one's self versus others; (b) assess outcomes that reflect cultural attitudes, mood changes, beliefs, physiological changes, and/or social effects; and (c) measure distinct versus

overlapping constructs. This variation has led some researchers to raise substantial concerns about measurement, in general, and construct validity, in particular. In addition, the nearly exclusive reliance on self-report questionnaires to measure expectancies is problematic to the extent that expectancies reflect cognitive processes that are nonconscious or automatic. Specifically, to the degree that one's beliefs about outcomes have at least a component that is reflexive, nonvolitional, and/or possibly not requiring attention or awareness, those beliefs cannot necessarily be captured by self-report questionnaires, which require deliberate introspection and awareness.

Despite measurement concerns, expectancies have been shown to be consistent predictors of behavior, especially alcohol consumption. For example, both positive and negative alcohol expectancies have been linked to alcohol consumption over time, including hazardous drinking, with some research suggesting that positive alcohol expectancies are better predictors. Responding to critiques in the field, researchers have continued to refine expectancy measures and analyses, and thus have been able to examine respondents' appraisals of specific outcomes (i.e. whether respondents view what researchers describe as "negative" outcomes as positive and vice versa). Similarly, the field has increasingly begun to identify and test moderators of expectancies and evaluate whether expectancies function as mediators of addictive behaviors. While much of the literature on expectancies in addictions has focused on alcohol and has relied heavily on college and adolescent samples, research has established that there are strong, positive relationships among expectancies and drinking behaviors. Less work has established the generalizability of these findings to other populations and/or other addictive behaviors. A growing literature has begun to show expectancy effects for marijuana, cocaine use, and gambling behavior. Additional studies – particularly those that establish causality, use prospective designs, and include diverse (clinical and nonclinical) populations – will be critically important.

MOTIVES

Whereas expectancies are presumed to precipitate behavior based on beliefs about probabilities of outcomes and subjective values associated with those outcomes, motives for engaging in addictive behaviors have been described as being more direct than expectancies. For example, an expectancy that alcohol will lead to tension reduction does not necessarily imply that a person will drink for this reason, even if the individual views tension reduction as a desirable outcome. Individuals typically have multiple motives for drinking or engaging in an addictive behavior. Therefore, on any given occasion, an individual may be drinking for social reasons even if he or she has positive tension reduction expectancies. Thus, motives are typically assessed by asking individuals about their reasons for engaging in addictive behaviors directly. The association between expectancies and behavior is likely to be mediated by motives, although some have argued that distinctions between expectancies and motives are primarily semantic.

Varying motives have been attributed to using different substances and behaviors that are typically classified as addictive. Although specific substances and behaviors are associated with unique motives, at least two categories of motives are common across substances and behaviors. These categories include social motives and affect regulation motives, which may be positive (e.g. affect enhancement) or negative (coping or tension reduction). These motives have been identified for drinking and marijuana use, but marijuana use is also motivated by unique factors, including "cognitive expansion" (e.g. perceptual and cognitive enhancement) and desires for experimentation. Similarly, gambling motives include not only social and affect regulation but also the desire to win money and the enjoyment of competition. Tobacco use also includes social and affect regulation motives. Notably, although common motives exist across substances and behaviors, their relative endorsement varies considerably. For example, stress reduction is a more commonly endorsed motive for cigarette smoking than for alcohol or marijuana use. Motives also vary in their relations to different outcomes. For example, social motives and enhancement motives are more strongly related to alcohol consumption, whereas coping motives are more strongly related to alcohol-related problems. Similarly, as alluded to above, motives are also likely to vary along the developmental course of an addiction, with curiosity, social facilitation, and positive affect enhancement likely to be more common in the initiation and early stages of the behavior. Engaging in addictive behaviors as a means of forgetting problems, regulating negative affect, and avoiding withdrawal symptoms are likely to be more prevalent later in the course of addiction.

Motives or reasons for changing behavior have also been extensively evaluated in the context of considering how, when, and why people successfully disengage from or quit addictive behaviors. In studies of natural recovery, former cigarette smokers and substance abusers who have successfully quit or learned to use moderately have been asked what prompted them to change. In these studies, health concerns are the most frequently endorsed motive for reducing or eliminating substance use. Financial reasons and negative personal reasons such as shame and guilt are also frequently endorsed reasons for change. A general assumption is that, over the course of addiction, individuals become

more motivated to quit or cut down as the perceived costs and negative consequences of the behavior begin to outweigh the perceived benefits and desired outcomes of the behavior.

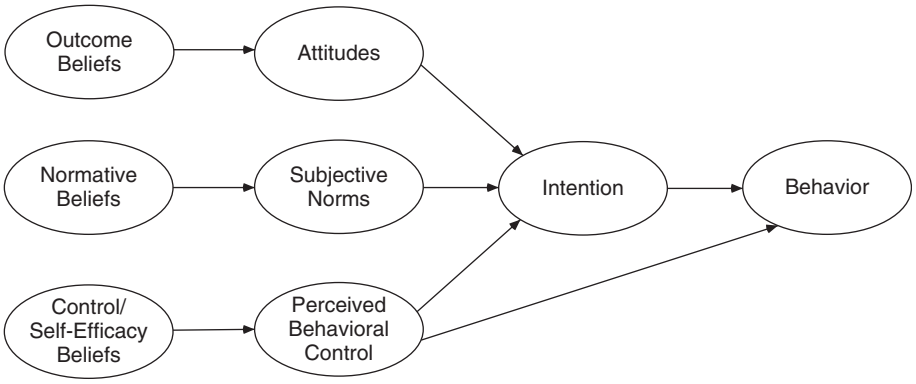
Identification of motives for initiation of use, continuing use, reduction, and abstention are critically important for the development of prevention and treatment strategies. Behavior change strategies should necessarily differ for an individual whose use is primarily motivated by perceived expectations from friends in comparison to an individual whose use is largely motivated around numbing negative emotions.

SOCIAL NORMS

Brief attention is provided here because the influence of social norms is inherently a cognitive process, and social norms are a consistent factor associated with initiation of substance use and other addictive behaviors. Many behaviors, including addictive behaviors, are influenced by other people. Specifically, our behaviors are influenced directly and indirectly by our parents, friends, peers, spouses, coworkers, children, and authorities, among others. Social norms refer to what is considered appropriate or normal behavior in a group of people. The influence that specific others have on us depends on how much we care about those others and our perceptions of their thoughts and actions regarding a specific behavior. One of the foundations of social psychology is the assumption that we are often not directly influenced by reality but rather by our subjective perception of reality. More importantly, in social contexts there are often discrepancies between subjective perception and reality. For example, substance use among adolescents and young adults is heavily influenced by perceptions of peer use and peer approval of use and it has been well documented that these perceptions are exaggerated. More generally, perceived social norms can increase or decrease the likelihood of engaging in addictive behaviors. We may correctly or incorrectly believe that our coworkers expect us to drink at the office Christmas party. One adolescent may smoke a cigarette if offered for fear of rejection, whereas another may resist when thinking about potential disappointment from parents.

BEHAVIORAL INTENTIONS

In contrast to our perceptions of others' expectations (i.e. social norms), cognitive influences on addictive processes also include our expectations of ourselves. Behavioral intentions refer to an individual's own expectations to engage in specific behaviors. Behavioral intentions are presumed to be proximal antecedents of behavior. Therefore, integrated models of addictive behaviors have sought to include proximal and distal predictors. Among the most widely studied of these models are the theory of reasoned action (TRA) and its extension, the theory of planned behavior (TPB). In these approaches to decision making, attitudes predict behavior through a single proximal antecedent, behavioral intentions. Intentions are typically defined as goals that are formulated through introspection or reasoning. Thus, from this perspective, the best predictor of substance use is the intention to use substances. Alternatively, the best predictor of changing use or entering treatment would be the intention to change or to seek treatment, respectively. Subjective norms and attitudes regarding the behavior influence one's intentions for engaging in the behavior. Subjective norms are one operationalization of social norms and refer to perceptions of the degree to which important others support the specific behavior (e.g. the extent to which friends and family approve of smoking cigarettes). Attitudes reflect the individual's global evaluations of performing a particular behavior (i.e. positive vs. negative evaluations of smoking). In general, the more favorable the attitude, the stronger the individual's intention is to perform it. The TPB is an extension of TRA that also includes measures of perceptions of factors that may facilitate or prevent engaging in a behavior (control beliefs), which impact perceived ease or difficulty of performing the behavior (perceived behavioral control).



The addition of control beliefs and perceived behavioral control provide additional information about potential constraints on action and were included to explain why intentions do not always predict behavior. For example, an individual may have favorable attitudes about substance abuse treatment and believe her family and friends are very supportive of her getting treatment. However, if she perceives that there are significant obstacles to getting treatment (e.g. that she has little control over being able to get treatment), she may be less likely to intend to and actually enter treatment. Each of the three major constructs (attitudes, subjective norms, and perceived behavioral control) is presumed to stem from cognitions that are behavior specific. For example, the TRA and TPB assume that decision making is a reasoned and deliberative process which involves consideration of behavioral options and anticipated outcomes. Two meta-analyses of TPB indicate that behavioral intentions explain 30–40% of the variance in health behaviors. The meta-analyses also found that TPB accounted for 27% and 39% of the variance in behavior and intention, respectively. The relative importance of attitudes, subjective norms, and perceived behavioral control in predicting intention is expected to vary across situations and behaviors.

Some researchers have argued that these rational models of decision making are less effective in predicting some kinds of behavior than others. Both TRA and TPB have proven effective at predicting health-promoting behaviors (e.g. family planning, cancer screening, exercise, treatment seeking) but have been less successful when applied to risky behaviors (e.g. heavy drinking and unprotected sex). This is especially true among adolescents as they are first exposed to risky health behaviors. The prototype willingness model (PWM) was created as an extension of TPB and, with the added component of behavioral willingness, explains how a common explanation for adolescent risky health behavior is simply that "it just happened." Willingness measures are designed to tap into an openness to engage in a behavior, or curiosity about a behavior that is internal rather than a reaction to social pressure. Behavioral willingness is reactive rather than deliberative, and that difference is the predominant distinction between behavioral intention and willingness. The PWM identifies intentions and willingness as independent predictors of behavior and posits that in the beginning stages of experience with the behavior, willingness is a better predictor than intentions.

SOCIAL COGNITIVE THEORY

Social cognition and social cognitive theories generally refer to our cognitions about others and about ourselves in relation to others. Within these domains, attributions are a central construct and relate to how we make sense of our own and others' behaviors, including addiction and substance/drug use. Attributions are explanations individuals give for the causes of events, others' behavior, or their own behavior. Attributions are made regarding why one's self or others misuse addictive substances, use moderately, or abstain altogether. People use different kinds of information when assigning causes to behaviors. For example, people consider consensus information or presumptions of how most people typically act in similar situations. Behavior that seems atypical is attributed to the individual versus to the specific circumstances of the situation. Thus, having a shot of vodka first thing in the morning is atypical, whereas a cup of coffee is not. A person who has a cup of coffee every morning will probably not be considered a caffeine addict but a person who has a shot of vodka every morning will probably be considered an alcoholic. Consistency information also plays a large role in attributions. Chronic use – meaning frequent, regular use – of any substance is almost a prerequisite for attributions of abuse or dependence. An individual who uses a substance once or only occasionally is unlikely to be considered addicted.

Individuals can make a variety of attributions for their own behaviors in the realm of substance use and abuse, and the nature of these attributions can have effects on subsequent behavior. For example, if a previously heavily drinking woman, who is a new mother, gets drunk after an extended period of abstinence, she could make a variety of attributions about her behavior, and these attributions can affect her subsequent use. Internal, stable, and global attributions, (e.g. "I just don't have much willpower," "I am a bad mother," or "I have a disease and therefore I simply can't control my actions"), in the absence of perceived alternative sources of available help, are likely to be associated with giving up or abandoning sobriety or attempts to moderate substance abuse. On the other hand, if she makes external, unstable, and local attributions, she may conclude such things as, "Everyone there was doing it, so it was a really difficult situation to be in and resist successfully," "I had the worst day of my life yesterday, so I was weak and turned to the alcohol/drugs to cope," or "I made a mistake, but this doesn't define who I am. I will forgive myself and get back on the path to abstinence." An event with attributions emphasizing situational (i.e. external) causes or exceptional factors will more likely be viewed as a mistake which can be learned from and allow one to go forth with new knowledge and be better prepared for similar future situations versus a certain verdict that one is "doomed" or "destined" to return to addiction.

In general, individuals tend to make trait or internal attributions when considering other people's behaviors and tend to make situational or external attributions when considering their own behavior. This is referred to as the fundamental attribution error or correspondence bias. For example, if you know that another person is getting drunk at a party instead of looking for a desperately needed job, you are more likely to attribute the cause of their behavior to something about their personality, and not the situation. However, when judging your own behavior, you bring self-knowledge and insight to the judgment, as well as greater knowledge of the environmental factors influencing you at the time. This results in a tendency to be more lenient in judgment. In the realm of addictive behaviors, this can lead to a tendency to generate more rationalizations for the self than an outside observer would, which can be either harmful or helpful in certain circumstances. If people make a mistake and have a small relapse, the tendency to be lenient in judging themselves can be helpful if they take from the experience knowledge of what led them to fail in that specific instance, as well as how they can better resist in the future. On the other hand, if people make mistakes and are lenient in self-judgment (i.e. continuously make excuses), but also never learn from their mistakes, then their leniency in judgment does not serve any constructive purpose in terms of improvement.

Another important cognitive factor related to regulating addictive behaviors is self-efficacy. Self-efficacy can be defined as individuals' beliefs in their personal capability to exercise control over a variety of tasks related to a specific domain. With respect to addiction, self-efficacy can be more narrowly defined as an individual's perceptions of his or her ability to control, moderate, or abstain from use of alcohol or other drugs. Self-efficacy largely determines an individual's performance in situations where there are opportunities to act upon personal goals. A person with high self-efficacy is more likely to press forward to accomplish their goals despite frustrations, physical or mental discomforts, or other challenges, whereas a person with low self-efficacy is more likely to abandon their goals when these challenges arise.

Self-efficacy plays a large role in determining performance; but it can also be very closely related to self-fulfilling prophecies. Self-fulfilling prophecies refer to instances where one holds a particular belief about oneself, acts in a way that validates the belief, and receives feedback from others that reinforces the belief. For example, if Susan thinks that she is too shy to make friends, Susan will act in ways that confirm her belief (e.g. by keeping quiet or by responding to friendly conversation with one-word answers). In turn, other people will likely respond to Susan's behavior by not initiating friendships with her. Susan, in turn, receives feedback from others that validates and strengthens her original belief about herself ("I am too shy to make friends"). The key to breaking the cycle of a self-fulfilling prophecy is to understand how one's own behavior (e.g. providing minimal responses to others who are trying to engage in friendly conversations) in response to a belief (e.g. "I am too shy to make friends") contributes to maintaining and confirming the belief (e.g. people may not have interest in pursuing friendships with someone who will not engage in conversation). Our understanding of attributions and self-efficacy are relevant in considering programs that emphasize complete abstinence. Specific consideration has been given to attributions people make when they violate abstinence goals. The same processes have also been discussed when people violate predefined limits.

ABSTINENCE VIOLATION EFFECT/LIMIT VIOLATION EFFECT

A specific process has been described regarding attributions that follow relapse after an extended period of abstinence or moderation. The abstinence violation effect can be defined as a tendency to continue to engage in a prohibited behavior following the violation of a personal goal to abstain. For example, an individual who has successfully abstained from alcohol, after having one beer, may drink an entire case of beer, thinking that since he or she has "fallen off the wagon," he or she might as well go the whole way. When an abstinence violation occurs, the attributions an individual makes play an important part in determining the trajectory of subsequent use. When abstinence violation occurs, individuals typically enter a state of cognitive dissonance, defined as an aversive experience resulting from the discrepancy created by having two or more simultaneous and inconsistent cognitions. Abstinence violators realize that their actions (e.g. "I drank") do not line up with their personal goal (e.g. "I want to abstain") and feel compelled to resolve the discrepancy. Attributions are made to try to resolve or justify the discrepancy. In this case, individuals try to explain to themselves why they violated their goal of abstinence. If the reason for the violation is attributed to internal, stable, and/or global factors, such as lack of willpower or possession of an underlying disease, then the individual is more likely to have a full-blown relapse after the initial violation occurs. On the other hand, if the reason for the violation is attributed to external, unstable, and/or local factors, such as an extremely tempting situation, then the individual is more likely to recover from the violation and get back onto the path of abstinence.

When abstinence is violated, individuals typically also have an emotional response consisting of guilt, shame, hopelessness, loss of control, and/or a sense of failure; they may use drugs or alcohol in an attempt to cope with the negative feelings that resulted from their abstinence violation. An abstinence violation can occur for a number of reasons. A person may experience a particularly stressful emotional event in their lives and may turn to alcohol and/or drugs to cope with these negative emotions. An abstinence violation can also occur in individuals with low self-efficacy, since they do not feel very confident in their ability to carry out their goal of abstinence.

These properties of the abstinence violation effect also apply to individuals who do not have a goal to abstain, but instead have a goal to restrict their use within certain self-determined limits. The limit violation effect describes what happens when these individuals fail to restrict their use within their predetermined limits and the subsequent effects of this failure. These individuals also experience negative emotions similar to those experienced by the abstinence violators and may also drink more to cope with these negative emotions. Cognitive dissonance also arises, and attributions are then made for the violation. In a similar fashion, the nature of these attributions determines whether the violation will lead to full-blown relapse.

IMPLICIT COGNITIVE FACTORS

Implicit cognitions are another important type of cognitive factor in addictions. They have received increasing attention in both the addictions and basic cognitive science literatures in the last 20 years as new techniques – often, but not exclusively, computer-based – developed to assess automatic cognitive processes. Implicit factors, also referred to as indirect and/or automatic factors, refer a class of factors that ostensibly occur without introspection, volition, attention and, possibly, awareness. They are thought to reflect cognitive processes that are more reflexive – that is, processes that are faster and simpler – as compared to those that are more reflective – that is, processes that are slower and more complex. Implicit factors are often categorized into one of three types: memory, attention, or interpretation factors. When individuals demonstrate a tendency or preference to associate one stimulus with another (e.g. associate alcohol pictures with good things vs. bad things); attend to a particular type of stimulus (e.g. an alcohol cue vs. a nonneutral cue); or interpret an ambiguous stimulus in a particular way (e.g. as an alcohol word vs. a neutral word), they are often described as demonstrating a bias (also called an information processing bias).

Implicit factors are often considered in tandem with explicit or controlled cognitive processes – that is, cognitive processes that appear to require introspection, volition, attention, and awareness. A variety of dual process models (i.e. models that focus on the influence of both implicit and explicit cognitive factors) have been proposed, including models specific to the development and maintenance of addictive behaviors. Common among dual models of addictive behaviors is the notion that addictive behaviors typically commence with explicit cognitive processes taking precedence. For example, a woman might decide to take drink because she is experiencing strong feelings of distress, or a man might decide to smoke a cigarette with friends because he needs a break from work-related duties. It is argued that, in both situations, the individuals are likely making conscious choices, selecting the substances from several deliberately considered options. Over time, it is argued that the "choice" of taking a substance becomes more automatic and less subject to introspection and that implicit factors may dominate or, at least, play a more influential role than explicit factors. For example, over time, the woman might attend more to cues related to alcohol when she experiences stress, and the man might tend to associate work breaks with cigarettes. Thus, repeated experiences are thought to lead to the development of biases related to implicit factors, which then, in turn, increase the likelihood of engaging in the addictive behavior. In addition, it has also been argued that specific circumstances or contexts increase the likelihood of implicit factors predominating over explicit factors, among them are experiencing strong emotions, being subject to a strong cognitive load, and being under the influence of substances. In such situations, effortful, deliberate processes are argued to be impaired whereas implicit factors are thought to be less affected.

As mentioned above, implicit factors are commonly, but not always, measured using computer-based tasks. Common computer-based tasks include adaptations of the implicit association test (IAT), single-target implicit association test (ST-IAT), extrinsic affective Simon task (EAST), dot-probe paradigm, the Stroop test, and lexical decision tests. Because implicit factors are thought to reflect more reflexive types of processes, these tasks usually include instructions to complete task trials as rapidly and accurately as possible. Thus, the primary analytical focus is on respondents' reaction time, with an occasional emphasis on respondents' error rates. Additional measures of implicit factors can be completed via pencil and paper and typically ask participants to respond to a word, phrase, and/or other stimulus as rapidly as possible. For those tasks, the content of the response is of primary concern. Within both the addictions and cognitive science field, there are ongoing

controversies regarding (a) exactly what implicit tasks are measuring; (b) the extent to which implicit tasks can capture "purely" automatic processes versus both automatic and controlled responses; and (c) the robustness (or lack thereof) of the psychometric properties of specific tasks. Such concerns are part of an even larger debate about what exactly is meant by the term "implicit" or "automatic" (i.e. those terms can be used to describe cognitive factors that do not require awareness, attention, volition, and/or consciousness) and the extent to which specific tasks can be said to measure all or some of those aspects is unclear.

Despite the substantial methodological concerns above, there is considerable promise with respect to implicit cognitive factors and addictions. The field is fairly young, yet despite its relative youth, it has already begun to address important questions and accumulate evidence that implicit factors are influential in addictions. In particular, research has shown that implicit cognitions related to alcohol, tobacco, and other drugs are associated with actual use and problems. Moreover, research has also shown that those cognitive factors predict variance in use and problems above and beyond explicit cognitive factors. Thus, there is evidence that implicit cognitive factors are uniquely associated with addictive behaviors. In addition, recent experimental studies have provided support that context – for example, exposure to alcohol cues, exposure to negative events, consumption of addictive substances, or depletion of self-regulatory resources – can affect implicit cognitive factors, and that in such circumstances, implicit cognitive factors may function as better, or at least unique, predictors of subsequent substance use than explicit factors. Similar to other types of cognitive factors, research on implicit factors is in need of more diverse samples (including clinical and nonclinical samples), prospective studies, and experimental research. Ongoing studies are also beginning to examine the extent to which implicit factors might also serve as additional targets for addictions treatment, which would further increase the potential importance of considering implicit cognitive factors in relation to addictions.

SELF-REGULATION

Daily, people successfully resist temptations such as overeating, excessive drinking, or ingesting harmful drugs through self-regulation. In sharp contrast, addictions and addictive disorders, including eating issues, substance abuse, and gambling, can be construed as examples of self-regulation failure. At the most basic level, self-regulation is the "exercise of control over oneself by oneself." Furthermore, self-regulation also includes the self's capacity to alter behaviors in order to achieve a desired goal state.

Self-regulation includes several theories, among them, and featured here, is Baumeister and Muraven's ego depletion theory, also known as "the strength model." This theory proposes that self-regulation depends on a finite resource of inner strength (ego) that becomes temporarily depleted after use. This depletion results in lowered ability to exercise self-control in subsequent tasks, even those under different self-control categories. Muraven and colleagues have likened ego depletion to a muscle – that is, a muscle has limited energy, and after doing a task, requires recovery time before it can perform another task at optimum levels. Similar to an athlete training muscles through repetitive practice, one can train the self-control "muscle," thereby improving self-regulation strength. Research has also shown that training can increase self-regulation in domains unrelated to the specific domain being trained. For example, participants who engaged in everyday self-control (e.g. improving posture) for two weeks performed significantly better in a dual-task paradigm (completing two unrelated self-control tasks consecutively) compared to a control group.

With regard to addictions, maintaining a balance between the desire to drink and the necessity of limiting intake requires self-regulation, and the strength model has been applied to addictions. For example, Muraven and colleagues found that compared to controls, participants who were instructed to resist thoughts about a white bear had lower self-control in a second task, namely resisting the temptation to drink in expectation of having to drive. Specifically, participants who focused on resisting thoughts consumed more beer and had higher blood alcohol content (BAC) levels than a control group. Research by Carey and colleagues has revealed that college students with better self-regulation skills reported fewer heavy drinking episodes and fewer alcohol-related issues. In that study, self-regulation was predictive of lowering both number of drinks consumed per day and peak BAC levels over 1 month following an intervention, thus, suggesting the potential utility of self-regulation training for preventing or treating addiction.

In addition to probable relationships between self-regulation and drinking, there also appear to be individual differences in self-regulation deficits, which are related to drinking behaviors. When moving toward addiction, the addictive behavior often becomes a habit (i.e. a behavior no longer under conscious control due to persistent repetition). As the habit continues to strengthen, the behavior becomes automatic and increasingly inaccessible to self-regulation. For example, some research has shown that self-regulation deficiency is positively related to habit, such that people with lower

amounts of self-regulation are more likely to develop habits, leaving them at an increased risk for developing an addiction. Thus, people with lowered self-regulatory resource levels can develop addictions because they are more likely to initiate alcohol consumption and fail to break their habit.

People may be unable to resist drugs or alcohol because of varying individual amounts of self-regulation and/or overuse of their self-regulation "muscle" in their lives. As the continued substance use progresses, behavioral habits can form, which may evolve into addiction. Repeated efforts that enhance the strength and endurance of self-regulation capacity may be effective in the long run. This may explain why repeated failures usually precede successful abstinence from smoking. Additional research is needed to further evaluate the efficacy of efforts designed to enhance self-regulatory capacity. As an alternative course, a number of studies have begun to explore strategies which functionally reduce cognitive effort expended in response to urges and cravings. A number of studies have evaluated mindfulness-based meditation, stress reduction, relapse prevention, and urge surfing as methods of reducing the cognitive effort expended in response to urges and cravings. The general idea here is to acknowledge desires or cravings without attempting to push them away, but simply observing them and watching them ebb and flow, without identifying with them or acting on them. Repeated practice of these approaches has shown some success with smoking and drinking. In summary, self-regulation is central to the construct of addiction and promising approaches for treatment from the perspectives of self-regulation theories are being investigated.

SPECIFIC SUBSTANCE EFFECTS

As described earlier, there are numerous cognitive factors that are influential in addiction. The aforementioned factors have applications to a range of addictive behaviors. There are also, however, some influential factors that are primarily associated with specific substances. Alcohol's effects on human social behaviors and emotions vary widely. For example, alcohol has been shown to elicit both aggression and altruism and both to reduce and increase anxiety. An obvious question is how can alcohol predict these contradictory behaviors? The notion that alcohol acts as a general disinhibitor and causes individuals to engage in risky and foolish behaviors is commonly accepted in the general public. Alcohol myopia theory was designed to provide a different perspective about how alcohol affects behavior and is one of the most accepted theories in this domain. It postulates that alcohol does not act as a general disinhibitor, but rather, that alcohol causes a restriction in controlled attentional capacity, such that intoxicated individuals no longer possess the executive cognitive processing skills to attend to all of the information in their environment. Disinhibition occurs, then, only when attention is focused toward salient instigatory (impelling) cues rather than inhibitory ones.

Sober individuals are able to attend to both impelling and inhibiting cues, evaluating the benefits and costs of engaging in various behaviors. For intoxicated individuals, complex deliberations about conflicts between competing impelling and inhibiting cues are precluded, and attention is granted only to the most salient and easily processed aspects of the situation. This disproportionate focus on the proximal environment represents a state of shortsightedness; the intoxicated person is only vaguely, if at all, concerned with considerations distal in time or place. In many situations, the impelling cues associated with risky behaviors (e.g. being sexually aroused) are more immediate than the inhibiting cues (e.g. the possibility of contracting an STD). In these cases, intoxication causes increases in reported intentions and performance of risky behaviors even when these behaviors contradict individuals' sober attitudes. Alcohol myopia theory is more comprehensive than disinhibition theory and explains a wide range of disinhibitory behaviors such as aggression, unprotected sex, driving under the influence, smoking, and suicide.

Alcohol is not the only drug that elicits specific cognitive effects. Caffeine, alcohol, and nicotine comprise the three most widely used psychoactive drugs in American society. Caffeine is an effective psychostimulant, ingested to obtain increased mental awareness, wakefulness, and a faster and clearer thought pattern. A consequence of this augmented mental alertness is continuous intellectual effort for prolonged periods of time without disruption of coordinated intellectual or motor activity. Caffeine and nicotine cause the least cognitive and behavioral impairment of all the addictive drugs. Nicotine is also a psychostimulant which results in increased cognitive functioning, attention, and memory consolidation. It can improve performance in a variety of cognitive tasks, the most consistent being on vigilance and rapid information processing. At higher doses, however, nicotine can induce nervous thought patterns which may develop into panic attacks.

Barbiturates are referred to as cognitive inhibitors because they are sedatives and depress memory functioning. The motor and cognitive inhibition and behavioral depression caused by barbiturates are similar to those caused by alcohol intoxication. In low doses, a person may respond either with a relief from anxiety or with withdrawal, emotional depression, or aggressive behavior. Higher doses elicit more general behavioral depression and sleep. One's mental state and social

setting interact to determine whether relief from anxiety, depression, aggression, or another unexpected response is experienced. Cognitive processing, judgment, memory, and motor skills become severely impaired during the period of intoxication, and may continue to be experienced for hours or days until the barbiturate is completely metabolized and eliminated.

Marijuana is the most commonly used illegal drug in the United States and is used for its relaxed, dream-like sense of euphoria. Two primary cognitive consequences of cannabis intoxication are impaired short-term memory and a perception that time passes more slowly. Marijuana also produces negative cognitive effects, such as distorted perceptions, disconnected or incoherent thoughts, and difficulty focusing, attending to, and solving problems.

Opiates (morphine, heroin, and numerous potent synthetic morphine-like compounds) rapidly produce an initial period of intense, euphoric, sublime pleasure. As the initial spike declines, a second phase begins – a pleasant, emotionally disconnected, dream-like perception, including strong feelings of contentment, well-being, and lack of concern. During this lightly sedated state, cognitive impairment occurs in the form of "mental clouding," accompanied by a lack of concentration, apathy, complacency, lethargy, and a sense of tranquility.

Psychedelic drugs (lysergic acid diethylamide (LSD), Phencyclidine (PCP), mescaline, and other hallucinogens) alter cortical functions, including cognition, perception, and mood. The psychological effects of LSD are alterations in perception, cognition, emotion, arousal, and self-image. Cognitive distortions include enhanced power to visualize previously seen or imagined objects and decreased vigilance and logical thought. PCP has been referred to as a "dissociative" drug, distorting perceptions and producing feelings of detachment. Users can experience several unpleasant psychological effects, with symptoms mimicking schizophrenia (e.g. delusions, hallucinations, disordered thinking, paranoia). Judgment and intellectual capacity are strongly impaired.

Cocaine and other amphetamines are powerful psychostimulants that markedly affect cognitive functioning and behavior. Temporary effects of cocaine in low doses include an immediate euphoria, giddiness, and enhanced self-consciousness and sense of grandeur. This period is followed by one of milder euphoria coupled with anxiety, followed by a lingering anxious state that spans hours. Thoughts typically race, and speech becomes extremely rapid, sometimes tangential or incoherent. Conscious awareness and mental acuity are increased but followed with a depression. As the dose or duration of cocaine use increases, all of these effects are intensified, followed by a rebound depression. Higher doses produce intense anxiety, hypervigilance, paranoia, and persecutory fears. These alterations in thinking patterns may lead individuals to become aggressive in response to imagined persecution. Other high-dose, long-term cognitive effects of cocaine include interpersonal conflicts, depression, dysphoria, anxiety, and vivid visual, auditory, and tactile hallucinations.

CONCLUSIONS

The purpose of this chapter is to provide a general overview of cognitive factors in addictive processes. As noted in the beginning of this chapter, cognitive factors comprise a wide range of more specific interrelated topics. In many cases, different terminology is used to describe fundamentally similar issues and ideas. For example, self-regulation is often referred to as willpower, and it seems clear that self-efficacy and attributions often lie at the heart of successes and failures of self-regulation. Moreover, for some researchers, the distinctions between motives and expectancies are only superficial. Regardless, both concepts seem to overlap considerably with descriptions provided in the TRA and the TPB. Social norms are also an inherent part of these theories as well as social cognitive theory and yet they are often considered independently from these theoretical confines. Across most constructs and perspectives is at least a background realization that much of what occurs is not entirely at the level of focused consciousness.

In considering cognitive factors involved in addictive behaviors, there are many perspectives which have evolved from traditions that vary in degree of interdependence. The variety of perspectives and overlap is directly related to training. Those trained in social cognitive theory are most likely to consider related topics from the perspective of social cognitive theory. Similarly, attribution theorists are likely to consider expectancies as specific types of attributions. One of the hopes for this chapter and the larger encyclopedia is to provide enough background so that readers can begin to consider connections and overlap across constructs in a way that will facilitate a deep and broad understanding of addictions. The specific content associated with this chapter is far from being a comprehensive survey of cognitive factors in addictive processes, but it is our hope that it provides a brief survey of the literature and a good introduction for the interested reader.

SEE ALSO

Emotions and Addictive Processes, Impulsivity, Disinhibition, and Risk Taking in Addiction, Peer Influences on Addiction, Substance Induced Myopia, Implicit and Associative Processes in Addiction, Craving and Expectancies

List of Abbreviations

TRA theory of reasoned action
TPB theory of planned behavior
PWM prototype willingness model

Glossary

Abstinence violation effect A tendency to continue to engage in a prohibited behavior following the violation of a personal goal to abstain.

Attributions Explanations individuals give for the causes of events, others' behavior, or their own behavior.

Behavioral willingness An openness to engage in a behavior, or curiosity about a behavior that is internal rather than a reaction to social pressure.

Controlled cognitive processes Cognitive processes that appear to require introspection, volition, attention, and awareness.

Limit violation effect A tendency to continue to engage in excessive behavior following the violation of a personal goal to moderate behavior.

Outcome expectancies An individual's expectations of the outcomes associated with drug use.

Perceived behavioral control Perceptions that one has control over his or her ability to perform a specific behavior.

Self-fulfilling prophecy A situation in which expectations create reality. For example, failure expectations may create reduced effort which, in turn, results in failure.

Subjective norms Perceptions of the degree to which important others support the specific behavior.

Further Reading

Ajzen, I., 1991. The theory of planned behavior. Organizational Behavior and Human Decision Processes 50, 179–211.

Carey, K.B., Henson, J.M., Carey, M.P., Maisto, S.A., 2007. Which heavy drinking college students benefit from a brief motivational intervention? Journal of Consulting and Clinical Psychology 75, 663–669.

Gerrard, M., Gibbons, F.X., Houlihan, A.E., Stock, M.L., Pomery, E.A., 2008. A dual-process approach to health risk decision making: the prototype willingness model. Developmental Review 28, 29–61.

Goldman, M.S., Del Boca, F., Darkes, J., 1999. Alcohol Expectancy Theory. In: Leonard, K.E., Blane, H.T. (Eds.), Psychological Theories of Drinking and Alcoholism, second ed.) Guilford, New York, pp. 203–246.

Jones, B., Corbin, W., Fromme, K., 2001a. Half full or half empty: the glass still does not satisfactorily quench the thirst for knowledge on alcohol expectancies as a mechanism of change. Addiction 96, 1672–1674.

Muraven, M., Collins, R.L., Nienhaus, K., 2002. Self-control and alcohol restraint: an initial application of the self-control strength model. Psychology of Addictive Behaviors 16, 113–120.

Sayette, M.A., 1999. Cognitive theory and research. In: Leonard, K.E., Blane, H.T. (Eds.), Psychological Theories of Drinking and Alcoholism, second ed.) Guilford, New York, pp. 203–246.

Wiers, R.W., Stacy, A.W., 2006. Handbook of Implicit Cognition and Addictions. Sage, Thousand Oaks, CA.

Relevant Websites

http://www.addictioninfo.org – Addiction Information.
http://www.people.umass.edu – Dr Ajzen's Theory of Planned Behavior.
http://www.niaaa.nih.gov – National Institute of Alcohol Abuse and Alcoholism.
http://www.nida.nih.gov – National Institute on Drug Abuse.

CHAPTER

28

Personality and Addiction Processes

Natalie Castellanos-Ryan, Patricia J. Conrod

Université de Montréal, Montreal, QC, Canada

OUTLINE

The emphasis placed on the role and significance of personality in the development and course of addiction has varied substantially through the past few centuries. While early models of addiction placed considerable emphasis on the role of personality, later more biological or pharmacological based models of addictions overlooked or disregarded personality as an etiological factor for addiction. During the 1960s and 1970s, new interest in the possibility of personality as a key factor in addictions arose, but was short-lived as, by the late 1970s, most of the studies carried out during this period failed to identify one single personality trait that conferred risk for addiction. In more recent years, however, models of addiction have placed a greater emphasis on the reciprocal interplay between biological, psychological, and environmental factors, and the heterogeneity of the addictive processes and profiles, identifying distinct pathways to addiction. One of the first and most influential of these models of addiction was proposed by Cloninger, who described two types of alcoholics. Type 1 was characterized by a relatively late onset, loss of control over drinking, experiencing guilt because of one's drinking, and high scores on the personality trait of harm avoidance (a trait associated with anxiety and pessimism). Type 2, on the other hand, was characterized by an early age of onset, displaying antisocial tendencies and behaviors, and high scores on novelty seeking (a trait related to impulsivity). Other influential studies carried out by researchers including Sher, Pihl, McGue, Stewart, Cooper, Conrod, and their colleagues, to name a few, have also highlighted the important role different personality and/or temperamental traits have in the development and maintenance of addiction. Among the personality factors most commonly cited in the literature as being associated with alcohol and drug misuse, as well as other addictions, are traits that fall within two main overarching personality domains of inhibited and disinhibited traits. Respectively and generally speaking, these two domains of personality correspond to a proneness toward negative and positive affective states, and the two main action tendencies of behavior, avoidance and approach, which are manifestations of the aversive and appetitive motivational tendencies in all humans. Within the inhibited domain, the personality dimensions that have been most consistently associated with addictive behaviors are (1) negative emotionality, introversion, or hopelessness; and (2) neuroticism, trait

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00028-0

anxiety, and anxiety sensitivity. Within the disinhibited personality domain, the two dimensions that have been most consistently implicated are (1) impulsivity, sometimes generally referred to as disinhibition; and (2) extraversion, including sociability and sensation seeking.

However, it is important to note that the association between personality and addiction is not simple or straightforward. Although these personality traits have consistently been linked with addiction, even across different forms of addictive behaviors, findings differ according to the methodology used and sample studied. For example, the personality traits that have often been found to characterize children of alcoholics or adolescent substance users are different from those that have been associated with clinical samples of substance misusers. Similarly, findings from cross-sectional studies highlighting concurrent associations between personality and addictions sometimes differ from the associations found in longitudinal studies. Differences could be attributed to the fact that the strength of associations between personality and addiction varies by developmental stage and, especially when it comes to substance misuse, at different stages in the course of the disorder, with severe or chronic substance misuse possibly resulting in changes in personality, negative affectivity, and impulsivity.

This chapter will provide a selective review of the literature establishing personality factors as correlates of, as well as risk factors for, addiction. Although addictive processes are similar across modalities or "objects of addiction," and similar personality traits have been implicated in a different range of addictive behaviors, it is important to mention that this chapter will mainly focus on findings related to addiction to substances such as alcohol and drugs. First, evidence linking specific personality traits to substance use and other addictive processes will be reviewed. The chapter will then review several causal pathways to addictive behaviors, in which personality traits play an important role. Finally, implications for prevention and clinical practice will be discussed.

PERSONALITY TRAITS IMPLICATED IN ADDICTION

Inhibited Domain

Research on the structure of neurotic symptoms and/or inhibited personality generally finds that negative affect is common to all neurotic traits but that they can be subdivided into two lower-order factors, characterized by low positive affect and fear, respectively. This hierarchical model of anxiety and depression also has relevance for understanding how neurotic/inhibited traits represent risk factors for addiction, but particularly for substance misuse. For example, while the higher-order factor of negative affect or neuroticism seems to be inconsistently related to risk for substance misuse, lower-order facets of neuroticism associated with low positive affect (e.g. hopelessness) and fear (e.g. anxiety sensitivity) have been consistently shown to have specific relationships to particular aspects of substance use and misuse.

Negative Affect and Hopelessness

Hopelessness has been defined, generally, as a set of negative expectations concerning self and future life, and, more specifically, as a tendency to expect negative things to occur, that nothing will improve negative circumstances, or, conversely, that highly desired outcomes will not occur.

In a number of cross-sectional studies, hopelessness or low positive affect has been associated with substance use and misuse. For example, hopelessness has been found to be associated with higher rates of alcohol abuse and dependence, sedative (including analgesics) drug use, and self-report reasons for substance use linked to depression coping and the numbing of painful memories in samples of college and secondary school students. Similarly, compared with adolescents with low levels of hopelessness, those with high levels of hopelessness have been shown to be at least twice as likely to report smoking tobacco in the last month, drinking in the last week, and using marijuana in the last month, as well as six times as likely to report cocaine use in the last month. Longitudinal studies confirm that hopelessness is a risk factor for substance misuse, showing that adolescents with high levels of hopelessness report greater acceleration in their trajectories of tobacco use, alcohol use, and marijuana use compared with those with low levels of hopelessness. These findings have now been replicated in American, Canadian, and British samples of adolescents. The association between this personality trait and addictive behaviors has also been found in clinical samples and community-recruited samples of adults. For example, Conrod and colleagues found in a community-recruited sample of alcohol-dependent women that a personality factor reflecting introversion and hopelessness was associated with a substance misuse profile that involved higher rates of dependence on analgesics and greater comorbidity with recurrent depression and social phobia. A study with college students also showed a greater likelihood of using sedatives/analgesics in college in those with increased hopelessness. The frequently reported finding linking hopelessness with a series of a number of internalizing problems, led researchers and clinicians in the

1980s and 1990s to hypothesize that hopelessness or low positive affect was associated with a particular susceptibility to substance misuse patterns through a self-medication process involving analgesia-induced numbing of painful experiences and memories. Through varied research, this hypothesis has been validated and has gained support as a specific pathway to addiction (*see* section "Casual Pathways to Addiction").

Anxiety Sensitivity

Another inhibited trait is anxiety sensitivity (AS), described as a fear of anxiety-related physical sensations due to an unrealistic expectation that they could lead to "catastrophic" consequences such as loss of physical or mental control, among others. Like hopelessness, AS has also been shown to be associated with high drinking levels and drinking problems, as well as with misuse of a variety of substances in adults, such as heroin, alcohol, nicotine, and anxiolytics. However, some of the findings linking AS and heavy alcohol use in community samples have been mixed and a number of studies have shown that AS is not associated with the use of marijuana, hashish, or stimulants. Other studies have shown increased levels of AS in individuals receiving treatment for substance use disorders, and Zvolenski and colleagues have shown that AS is associated with heightened reactivity to nicotine withdrawal and the tendency to rapidly return to smoking during quit attempts.

Although there is some prospective evidence that anxiety symptoms in childhood or adolescence often precede substance use and misuse, very few prospective studies on the association between AS and addiction have been undertaken. Of those that have, findings are mixed or suggest some developmental and gender specificity. For example, one study carried out in Finland found that anxiety/shyness assessed at the age of 8 predicted increased alcohol and other drug use 20 years later in women, but predicted reduced substance use in men. In contrast, another study carried out in New Zealand by Caspi, Moffitt, and colleagues showed that inhibited boys, described as fearful, anxious, and shy, were more likely to present alcohol-related problems. However, studies assessing AS in early adolescence and its association with the onset of or increased substance use across adolescence have not been able to confirm a prospective relationship between AS and early onset or higher severity of substance use across this period. However, the prospective link between AS and development of alcohol use disorders, becomes clearer in young adulthood, with studies carried out by Schmidt and colleagues, as well as Stewart and colleagues, showing that AS assessed at ages 18–20 predict alcohol use disorders a couple of years later. A possible reason for the inconsistencies in this literature is offered by O'Connor and colleagues who have shown that alcohol and drug expectancies, especially those related to tension reduction and cognitive and behavioral impairment, are important moderators in the association between AS and substance use but are not often taken into account.

What has been consistently reported in the literature, in both adolescent and adult samples, is the significant association between AS and self-report motivations for substance use that reflect self-medication of anxiety symptoms or drinking to cope with emotional distress. Many have also demonstrated that AS is associated with a pharmacological sensitivity to the arousal dampening properties of alcohol and benzodiazepines. Indeed, as AS has often been conceptualized as an "arousal-accelerator," it is not surprising that the evidence does not support AS as a risk factor for adolescent onset substance use, but rather indicates that AS represents a specific risk profile predicting who will use substances to cope with stressors that produce physiologic arousal, such as drug withdrawal, trauma, severe negative life circumstances, and, particularly in younger samples, those normative experiences such as social pressure to conform to peers' substance use patterns. It is important to note here, that there is some literature suggesting that high AS men are at greater risk for self-medication drinking than high AS women, especially when tension reduction expectancies are increased, and that, in more general terms, drinking for coping with negative affect (also related to hopelessness) is a stronger motivation for men than for women.

Disinhibited Domain

Disinhibition is often referred to a general failure to plan, control, or regulate behavior, especially behavior that can be unduly risky or can sometimes result in negative consequences. Recent research on the structure of disinhibited behavior and personality also supports a hierarchical structure of disinhibited or externalizing disorder symptoms and traits, indicating that while they are all correlated with each other and share common variance, subfactors exist that capture the unique variance of some disinhibited behaviors and traits. Certainly, disinhibited tendencies have been referred to in a number of ways in the personality, behavioral and psychopathology literatures, from "acting without premeditation," "lack of planning," "excitement seeking," "low tolerance to boredom," "behavioral undercontrol," and impulsivity, among others. Consequently, measures labeled "disinhibition" or "impulsivity" may measure different constructs from each other. Indeed, although general factor models of personality typically identify only one factor for

disinhibition/impulsivity, extensive research has been carried out on the differentiation between different dimensions of disinhibited personality, which usually results in between two and four subfactors of disinhibition. Using methodologies like factor analysis, recent research in the field of personality has identified, as many as four personality facets associated with impulsive-like behavior: lack of planning, lack of persistence, urgency (acting rashly when upset or anxious), and sensation seeking, but most studies in the field of personality as well as cognitive sciences, agree that at least two clear subdimensions of disinhibition exist: one which is referred to as impulsivity in this chapter, and another which is referred to as sensation seeking.

Impulsivity

Impulsivity is generally associated with a deficit in reflectiveness and planning, rapid decision-making and action, and a failure to inhibit a behavior that is likely to result in negative consequences. It is clear from the literature on substance misuse in adolescent and adult samples that impulsive traits play a prominent role in addictive behavior. Impulsivity has often been associated with substance misuse, specifically, quantity and frequency of drug use, and early experimentation with drugs. Impulsivity has also been associated with the consumption of a range of different drugs, such as cannabis, ecstasy, and heroin use, and is the personality trait that has most consistently been associated with alcohol disorders in the literature.

Longitudinal studies have also identified impulsivity/disinhibited traits as risk factors for future substance misuse. For example, impulsivity in childhood, as measured by either Eysenck's psychoticism or Cloninger's novelty seeking, has been found to predict substance misuse and alcohol-related problems in adolescence and adulthood, with a study by Cloninger and colleagues showing those reporting high scores in impulsivity to be 20 times more likely to report alcoholism in adulthood compared with those who do not report high scores on this trait. Impulsivity measured by higher-order factors or super-factors like constraint (i.e. Multidimensional Personality Questionnaire or the Minnesota Multiphasic Personality Questionnaire) or by lower-order, more specific measures of impulsivity (e.g. Substance Use Risk Profile Scale) assessed in childhood or early adolescence, have also been prospectively associated with alcohol and drug use in late adolescence and young adulthood. However, it is important to highlight that some laboratory studies have shown that severe and persistent substance use can result in deficits in behavioral and/or cognitive measures of impulsivity, such as response inhibition and decision-making, and even in increased levels of self-report trait impulsivity, suggesting the possibility that the pathway from impulsivity to substance misuse is bidirectional. This highlights the importance of assessing both measures of impulsivity and substance use simultaneously across time, so that cross-lag or bidirectional effects can be evaluated. With this in mind, a recent study by Littlefield, Sher, and Wood evaluated, using latent growth models, the extent to which changes in personality and changes in drinking behavior covary across early adulthood and showed that changes in drinking behavior from 18 to 35 years of age tend to co-occur with changes in impulsivity (as well as neuroticism). While this analysis does not provide direct evidence for causal effects between these two factors, the findings suggest that the relationship between impulsivity and drinking behavior may be more complex than a simple causal relation and may be reflecting a mutually exacerbating relationship.

There is now enough evidence – provided particularly by studies on children of alcoholics as well as by studies on nonsubstance use–related addictions such as problem gambling, where the vulnerability mechanisms underlying addiction can be studied without the confounding effects of substance use – showing that impulsivity is indeed a risk factor for addiction. For example, studies on children of alcoholics, considered to be at a higher risk for future alcohol problems due to genetic vulnerability, have found that many of these children exhibit high levels of impulsive personality traits and behaviors, with some even showing that disinhibited traits mediate the relationship between family history of alcoholism and drinking behavior.

Studies on problem gambling have shown similar results than those in the substance use field, indicating that compared with nongambling controls, problem gamblers demonstrate increased scores on a range of impulsivity-related measures. Similar to findings in the alcohol, and drug use literature, impulsivity scores have also been shown to predict gambling symptom severity in clinical samples of problem gamblers, as well as community-recruited adolescents. Finally, several studies provide evidence for the hypothesis that impulsivity is a common risk factor for (or associated with shared vulnerability to) gambling and substance use problems, as well as other addictive behaviors and externalizing problems, such as antisocial behavior.

Although there is some research showing that impulsivity (or undercontrol) assessed early in childhood is associated with alcohol problems in early adulthood in men but not women, varied research has shown that men and women tend to report similar levels of impulsivity – this is not the case though if the impulsivity measure includes "aggressive tendencies" (such as the measure of constraint) on which men score higher – and similar associations between

impulsivity and later substance use or other addictive behaviors.

Sensation Seeking

Sensation seeking is generally defined as a strong need for stimulation, a low tolerance to boredom, and a willingness to take risks for the sake of having novel and varied experiences. Findings on the association between traits related to extraversion and substance misuse in clinical samples as well as community samples of adults have been somewhat inconsistent, with a number of studies indicating only modest associations between extraversion and drinking onset as well as increased levels of alcohol use. However, sensation seeking, a trait related but not equivalent to extraversion, has been shown to be more robustly and consistently related to substance misuse behaviors, especially binge drinking or heavy episodic drinking, particularly in adolescents and young adults. Indeed, a recent study has shown that in a sample of drinking college students, sensation seeking explained a significant amount of variance in alcohol-dependence symptoms, above and beyond that explained by the trait of extraversion as measured by the Neuroticism Extroversion Openness Five Factor Inventory (NEO-FFI). Research carried out by Cooper and colleagues and Stewart and colleagues have consistently shown that sensation seeking is associated with using substances for enhancement motives rather than for social or negatively reinforcing motives.

Longitudinal studies in North American as well as European samples have shown that sensation seeking is not only associated cross-sectionally with binge drinking but can also predict drinking rates and growth in drinking quantity by frequency during adolescence. Although findings are not as consistent as with alcohol use, some studies have shown prospective association between this trait and marijuana, tobacco, and hallucinogen use in adolescence. Most of the positive findings linking sensation seeking and future substance use have been found in adolescent or young adult samples, suggesting that sensation seeking might be an important factor in earlier stages of substance use initiation or experimentation, rather than in the continuation or compulsive use.

Although men tend to score higher on sensation seeking measures than women (especially during adolescence), most research shows that the association between high Sensation Seeking and substance use is significant for both men and women, with only a few studies suggesting that the association between sensation seeking and alcohol use outcomes may be stronger for adolescent boys than girls.

A "need for intensity" and sensation seeking have also been associated with problem gambling. Some studies have characterized pathological gamblers as having a low tolerance to boredom, being attracted to stimulating situations, and finding it difficult to complete tasks they find boring – all characteristics used to describe sensation seekers. However, although often hypothesized that by their very nature sensation seekers should be more prone to gamble, findings have been mixed. While several studies have reported high levels of sensation seeking in adult pathological gamblers, particularly those who bet at race-tracks or casinos, this has not been the case in younger problem gamblers or across all gambling activities. In fact, one study by Coventry and Brown has found that those who bet exclusively in an off-course betting office scored lower on sensation seeking than non-gamblers and the general population. Findings like these highlight the fact that gambling, and addiction generally, can't be viewed as a homogeneous activity or process and that different pathways to addiction as well as different profiles of addiction can be identified.

CAUSAL PATHWAYS TO ADDICTION

The evidence reviewed above implicating personality as an important factor in addiction is but a few of the several longitudinal studies, familial aggregation studies, as well as genetic and animal studies, which provide strong evidence for causal relationships between personality and addictive behaviors. In line with the four personality traits described above, the available evidence suggests at least four different developmental or causal pathways to addiction, in which the different personality factors play an important etiological role. It is proposed that these causal pathways are described as (1) the negative affect regulation pathway, (2) the stress-dampening pathway, (3) the poor behavioral inhibition and/or deviance proneness pathway, and (4) the reward sensitivity pathway.

The Negative Affect Regulation Pathway

Several studies have shown that many individuals engage in heavy drinking or substance use, as well as excessive gambling or eating, to regulate affect or emotional states. Also referred to as the "self-medication" hypothesis, many individuals report that they use substances to regulate or cope with negative affect, such as depressed mood or anxiety, and to forget about difficult situations. Although the negative affect regulation model is one of the most enduring etiological perspectives on addiction, it has been shown that it is highly dependent on intraindividual factors such as personality, as well as other factors such as expectancies, genetics, and environmental factors.

Inhibited traits have typically been implicated in models of negative affect regulation of addictive

behaviors. For example, hopelessness has been argued to reflect sensitivity to punishment and has been linked, by researchers including Sher, Cooper, and colleagues, to the development of alcohol use by using its analgesic properties to suppress feelings of negative affect. Fitting with this profile of addiction are numerous findings indicating that depression and negative affect and/or hopelessness generally precede addictive behaviors (although it is important to note that extended substance use and gambling can also result in increased depression or feelings of hopelessness), and recent evidence that the link between hopelessness and future substance use is mediated through the occurrence of depressive symptoms, and motivations for drinking that include coping with depression.

Anxiety sensitivity has also been associated with substance misuse for negative affect regulation, but specifically to relieve feelings of anxiety. Consistent with this are studies by Kushner and colleagues and Stewart and colleagues showing that AS is associated generally with coping motives for alcohol use, and studies showing that those high in AS use substances specifically to avoid or escape anxiety symptoms.

Accordingly, this pathway to addiction has been associated with late onset, a greater tendency to engage in addictive behaviors (i.e. use substances or gamble) in solitary contexts, and, specific to substance use, a predilection for alcohol, sedatives, and/or analgesics, as well as reporting lower enhancement motives for use.

The Stress-Dampening Pathway

Over the last 20 years, a large body of research has assessed individual differences in their response to the effects of alcohol and other drugs, which may put certain individuals at higher risk for addiction. First proposed by Sher and colleagues, the pharmacological vulnerability model of addiction proposes that individuals are at risk for substance misuse because they are especially sensitive to the reinforcing effects of substances and are therefore more likely to use substances as they experience greater effects from the substance.

Inhibited traits such as AS and low positive affect have been implicated in differential psychopharmacological effects of substances. AS has been associated with experiencing increased withdrawal symptoms, particularly those related to tobacco, and thus poorer cessation outcomes. Similarly, low positive affect has been associated, even after controlling for depression symptoms, with higher withdrawal symptoms in a sample of adults attending a smoking cessation clinic. However, studies specifically assessing the stress response–dampening effects of alcohol, have implicated AS as a determining factor. Often described as an arousal-accelerating factor, AS seems to not only render individuals susceptible to high levels of arousal in normal stressful situations, but also in response to acute and chronic drug withdrawal, physiologic states that are easily dampened by the pharmacologic properties of alcohol and benzodiazepines.

Several studies have found that individuals with a family history of alcoholism often have an increased sensitivity to the dampening effects of alcohol on stress response, with findings showing that personality factors related to anxiety and AS seem to play an important role. For example, individuals who are high on inhibited traits like AS have been shown to display reduced electrodermal activity to threat cues when moderate to high levels of alcohol have been consumed. Similarly, Conrod and colleagues found that men with higher self-reported levels of anxiety sensitivity experienced electrodermal response and heart rate dampening effects to aversive stimulation after alcohol administration, compared with low anxiety sensitive men. These findings were interpreted as showing a pharmacologic sensitivity that is produced by an interaction between anxious personality and sedative drug effects to produce a highly negatively reinforcing fear reduction. This is further confirmed by findings from O'Connor and colleagues showing that risk for alcohol misuse in men results from increased scores on AS in combination with the belief that drinking alcohol can reduce tension (i.e. tension reduction alcohol expectancies). This combination of factors – high AS, the belief that alcohol or drugs can reduce tension, and the heightened sensitivity to their pharmacological effects – together with the fact that chronic substance use will, with time, dampen or reduce the "tension-reduction" properties of the substance being used, leading to further substance use, might explain why this pathway is associated with substance use dependence or problems, rather than with substance use onset.

The Poor Response Inhibition Pathway

This pathway to addiction posits that individuals with high levels of impulsivity (or low levels of constraint) are more inclined to engage in deviant behaviors in general, including substance misuse and other addictive behaviors. This pathway to addiction has been frequently investigated and thus is well documented. There have been a few models that have attempted to explain the link between impulsivity and/or poor response inhibition and deviant behaviors as well as substance misuse. As early as the 1970s, Jessor and Jessor, and other colleagues, believed that it was the interplay and interdependence of three systems of variables – the behavior system, the personality system, and the perceived environmental system – which resulted in the engagement in problem behaviors or level of

behavioral deviance, of which substance use was considered only an indicator. Although emphasis is placed on deficient socialization as a major "instigator" or risk factor, their model highlighted the role temperament and personality traits, particularly those related to impulsivity, had in socialization and developmental processes. Two more recent models explicitly implicate impulsivity or deficits in self-regulation as key in the development of behavioral problems and substance misuse, as well as their comorbidity: the social deviance model, proposed by Sher and colleagues, and the psychological dysregulation theory, proposed by Tarter, Clark, and colleagues. Both highlight the interplay between individual and environmental factors, and refer to substance use as part of a more general deviant pattern of behavior, which usually begins in childhood, and can be exacerbated by poor socialization. However, while the deviance proneness model is firmly based on sociopsychological theory, and places greater emphasis on the role of poor socialization, the psychological dysregulation theory is based on psychobiological theory of human behavior, and places greater emphasis on genetic liability. Consistent with both these models, longitudinal studies assessing early onset alcohol problems show consistent associations between alcohol problems and a history of childhood antisocial behavior, poor school achievement, poor interpersonal relationships, heightened activity or attentional problems during childhood, and inadequate parenting. Furthermore, studies have shown that early "difficult temperament" characterized by high levels of disinhibition or impulsivity, in combination with poor parenting, leads to unsocialized behavior. Impulsivity has also been shown to be associated with substance misuse that is comorbid with antisocial behavior, while other disinhibited traits like sensation seeking have not. Recent findings, using structural equation modeling which make it easier to assess the comorbidity between disorders, as well as their longitudinal associations, have also provided support for this pathway to substance misuse. For example, one study demonstrated that adolescents high in impulsivity showed their susceptibility to increased alcohol use through conduct disorder symptoms, whereas adolescents high in sensation seeking showed a direct susceptibility to increased alcohol use and were only susceptible to conduct disorder symptoms as a consequence of their increased alcohol use. Furthermore, another study by Castellanos-Ryan and Conrod showed that while impulsivity was associated with an externalizing behavior factor, which accounted for the shared variance between measures of conduct disorder and substance misuse, as well as a specific conduct disorder factor, it was not associated with substance misuse that did not co-occur with other externalizing behavior problems in adolescence. Finally, a study by Khan and colleagues investigating the structure of psychiatric disorders in adults also showed that impulsivity (i.e. novelty seeking), but no other personality traits, accounted for a proportion of the comorbidity between dependence to substances (alcohol and drug) and conduct disorder. These results, and the high rate of co-occurrence between substance misuse and antisocial behaviors, lend support to the deviance proneness and the behavioral dysregulation models of substance misuse and also suggest that impulsivity, but not other personality traits, may play a key role in these liability models.

Recent studies which use behavioral or cognitive measures of inhibition, also provide support for this pathway to addiction, with findings consistently showing that deficits in self-regulation, assessed by increased commission errors on go/no-go or stop tasks, are associated with a range of behavioral problems, including conduct disorder, hyperactivity, and substance misuse. Moreover, Castellanos-Ryan and Conrod showed that poor response inhibition, as measured by increased commission errors in a STOP task, partially mediates the association between self-reported impulsivity and externalizing behaviors, including substance misuse and conduct disorder symptoms. These results indicate that the link between impulsivity and substance misuse is, at least in part, explained by a deficit in response inhibition that makes impulsive individuals more prone to engage in externalizing behavior in general.

Finally, further support for this pathway and a common genetic liability for deviant or externalizing behaviors comes from several studies showing common genetic liability for conduct disorder, substance use, behavioral undercontrol, as well as studies by Iacono, McGue, and colleagues, as well as Slutske and colleagues showing common genetic liability between early measures of social deviance and later externalizing disorders and substance misuse. Although it seems likely that a large number of genetic variants are implicated in externalizing and substance use behaviors, several common variants have been implicated in these behaviors and their comorbidity, with some studies suggesting that a significant portion of the genetic contribution to early onset problem drinking and other drug use is mediated by personality. For example, studies have shown that the serotonin transporter (5-HTT) and the low variant of the monoamine oxidase A (MAOA) are implicated in substance use disorders and other problem behaviors, as well as being associated with the personality trait of impulsivity and, in the case of MAOA, with neurocognitive measures of disinhibition. Other genes, such as the serotonin transporter polymorphism (5-HTTLPR), the DRD2 polymorphism, and the D4 dopamine receptor (DRD4) have been shown to be

associated with impulsivity or disinhibited personality and substance misuse.

Besides the high rates of co-occurrence or comorbidity with other problem behaviors, such as antisocial traits and behavioral problems as mentioned above, this pathway to addiction has been characterized by early onset, more severe levels of substance used (as well as of gambling), as well as undefined motives of substance use and a lack of specificity to substance or addictive behavior of choice (i.e. impulsivity has been implicated in a wide range of addictive behaviors and substances).

The Reward (and Positive Reinforcing Properties of Substances) Pathway

This pathway to addiction posits that those individuals scoring high on sensation seeking or related traits, such as novelty seeking, extraversion, or reward seeking, will engage in addictive behaviors motivated by their positive reinforcing properties or to enhance positive feelings or emotions. Supporting this pathway to addiction are several studies, including those by Pihl, Peterson, and colleagues, that show that heightened physiological sensitivity to substances, such as accelerated heart rate after alcohol consumption, is associated with increased risk for alcoholism and substance misuse. Similarly, compared with controls, men with a family history of alcoholism have been shown to experience a heightened subjective stimulant response to amphetamine. In another study, college students with poor inhibitory control reported heavier drinking and exhibited enhanced subjective stimulation during the ascending limb of the blood alcohol curve. However, although disinhibited traits like impulsivity and sensation seeking have both been associated with increased and problematic substance use, only sensation seeking has consistently been associated with enhancement or reward-related motives for substance misuse. The motivation for positive reinforcement from substances has been shown to be founded on the neuropharmacological effect substances have on the brain centers involved in basic reward mechanisms, that is, substances stimulate mesolimbic dopamine activity and increase activity in brain opioid systems. Consistent with this, are studies showing that individuals who are high on disinhibited personality traits such as sensation seeking tend to be more sensitive to drug-induced reward and to display heightened heart rate response to alcohol, as well as experience more positive feelings in general after alcohol intoxication. Thus, the link between sensation seeking and substance use could be partly explained by the association between sensation seeking and this heightened response to the reinforcing properties of substances. Another possible explanation of the association between sensation seeking and substance misuse is offered by Leyton and colleagues, who found, in an exploratory study using positron emission tomography (PET), that sensation seeking (as measured by "exploratory-excitability," a subdimension of Cloninger's novelty seeking) was associated with greater amphetamine-induced dopamine release in the ventral striatum and drug wanting. The authors suggested that amphetamine consumption elicits a dopamine-mediated appetitive state which is stronger for those high in sensation seeking.

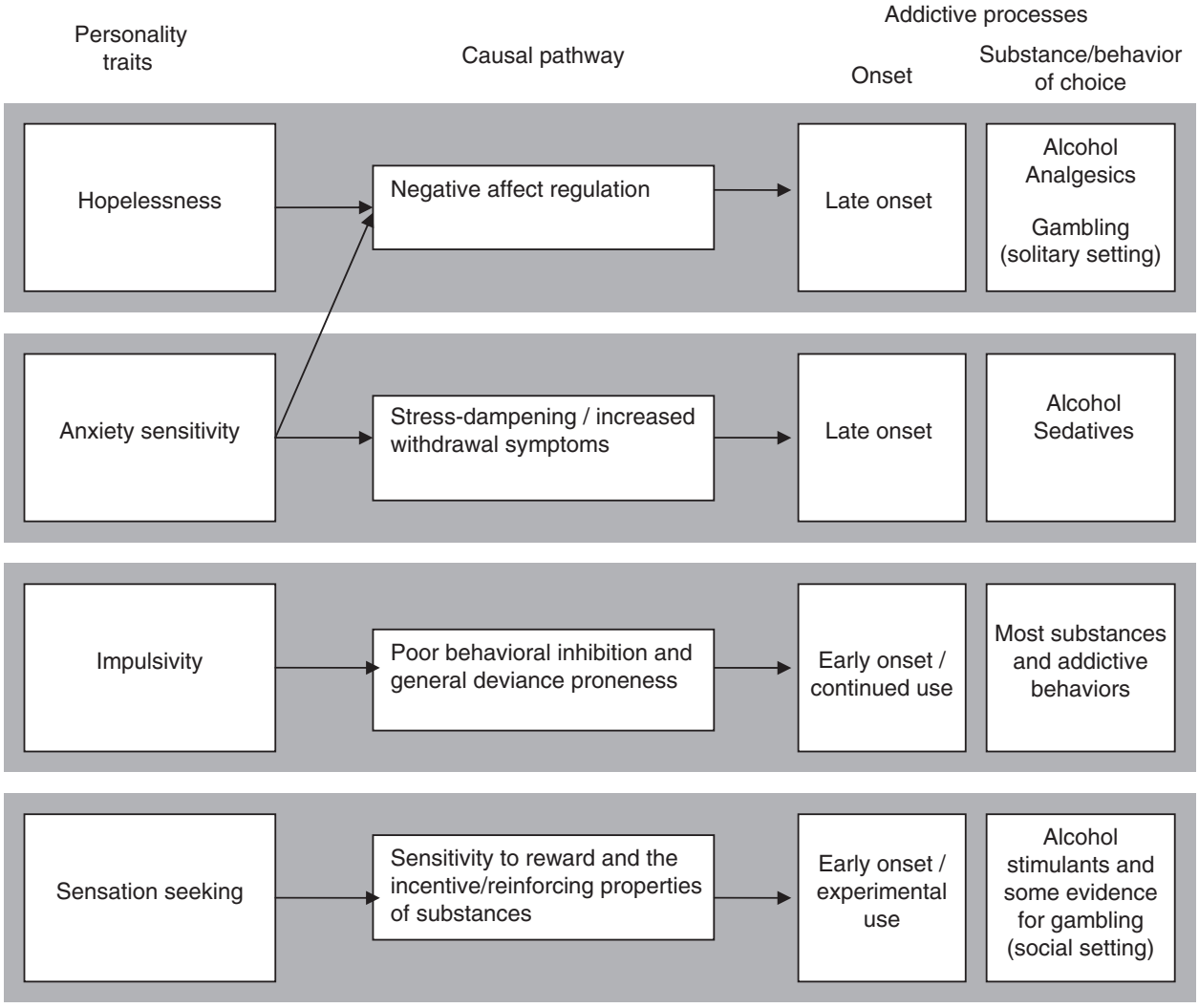
However, it is important to note that sensation seeking has also been linked to risk-taking in general. Accordingly, sensation seeking has been associated with general or nonsubstance-related measures of reward sensitivity/dependence, particularly in studies looking at gambling behavior. Moreover, a recent study by Castellanos-Ryan and Conrod has shown that the association between sensation seeking and binge drinking in adolescence was partially mediated by reward sensitivity, as measured by a (monetary) rewarded go-no-go task, while other measures of disinhibition, such as impulsivity, were not.

SUMMARY AND IMPLICATIONS

The development of addictive behaviors seems to be multidetermined and multiple risk factors for addiction have been identified, such as genes, individual differences in sensitivity to the reinforcing effects of substances, age of onset, and the presence of deviant peers, as well as conduct and emotional problems. However, it is also clear that some individuals, sometimes subjected to the same general risk factors, go on to develop an addiction and some do not. As reviewed, a number of research findings suggest that different pathways to addictions exist, where psychological factors such as personality traits, play an important role. These different casual pathways to addiction suggest that personality traits are related to substance misuse, and possibly other addictive behaviors, through different motivational processes, and are associated with different profiles and/or patterns of addiction. For example, while findings support that anxiety sensitivity is associated with substance misuse through its association with the anxiolytic or stress-dampening effects of specific substances, sensation seeking is associated with the initiation of substance use behaviors as well as substance misuse through the mediation of reward sensitivity and the positive reinforcing effects of certain substances. On the other hand, impulsivity seems to be associated with a large number of addictive behaviors and processes through a general inability to inhibit responses or behavior, while findings seem to

suggest that hopelessness is associated with substance misuse through a motivation to cope with high levels of negative affect.

Displayed graphically in Fig. 28.1 are the four hypothesized causal pathways to addiction reviewed in this chapter. Described in this figure are the distinct personality traits associated with each pathway, as well as the underlying mechanisms that help explain the way in which these personality traits are related to addictive behaviors. Included in Fig. 28.1 are also what one could refer to as addiction profiles – e.g. early or late onset, and specific substances or addictive behaviors – most commonly associated with each of the four pathways. The figure shows how individuals high in hopelessness are likely to engage in addictive behaviors through their need to manage or reduce their negative affect, while individuals high in anxiety sensitivity engage in addictive behaviors, primarily substance use, because of their sensitivity to the arousal dampening effects that are very negatively reinforcing to those who fear them. Both these pathways to addiction are associated with later onset, are more likely to lead to gambling or using



Note: This figure shows that individuals high in hopelessness engage in addictive behaviors that usually have a late onset through their need to regulate their negative affect. Their drugs of choice are usually alcohol or any other with analgesic properties. Individuals high in anxiety sensitivity also engage in late onset addictive behaviors, such as misuse of alcohol or substances with sedating properties, because of their sensitivity to their arousal dampening effects. Those high on the trait of impulsivity tend to engage in a range of addictive behaviors and have a more severe addiction profile, through deficits in response inhibition. Individuals high in sensation seeking, due to their specific sensitivity to reward, including alcohol- or drug-induced reward, are likely to engage in early onset or experimental substance use, especially with those with psycho-stimulant properties, or gambling in social settings.

FIGURE 28.1 Personality-related causal pathways to addiction.

substances in solitary contexts, and opting for substances with analgesic effects, in the case of hopelessness, or anxiolytic effects, in the case of anxiety sensitivity.

With regard to the more disinhibited pathways to addiction, Fig. 28.1 shows a third pathway indicating that those high on the trait of impulsivity are likely to engage in a range of addictive behaviors, as well as to engage in continued or severe addiction, through deficits in response inhibition. The fourth pathway illustrates how individuals high in sensation seeking through their specific sensitivity to reward, including alcohol- or drug-induced reward, are likely to engage in early onset or experimental substance use, develop drug-taking patterns to enhance psychostimulation (e.g. binge drinking), and/or engage in gambling, particularly in casinos or race-tracks or other social contexts, where emotions related to gains and losses are heightened. Although these two pathways are similar in that they are both related to early onset of addictive behaviors, they differ in a number of ways, one of the most important being that individuals high in impulsivity, as opposed to high sensation-seekers, in the context of substances that produce potent psychostimulant reward, are not protected by an intact response inhibition system and more easily develop compulsive (or uncontrolled) self-administration.

It is important to highlight that these pathways to addiction, like personality dimensions, which should not be considered as discrete but continuous measures, are not mutually exclusive and can overlap in some individuals. Thus, for example, individuals might experiment with alcohol or other substances in adolescence to satisfy their need for rewarding, novel or intense experiences, but might then be unable to control their substance use due to their impulsivity, and thus, their inability to self-regulated behavior.

Although addiction is a complex phenomenon, involving a range of biological, environmental, and psychological factors, the evidence reviewed identifies personality traits as important etiologic and moderating factors. This evidence makes it clear that addiction is not homogeneous and thus highlights the need to either tailor intervention approaches to specific profiles of addictions or to target personality profiles directly along with treatment as usual. The evidence also identifies personality traits or the different pathways to addiction as key targets for prevention efforts. Although most current prevention approaches target behaviors, such as substance use or gambling, directly, it seems more logical that interventions would want to target liability factors rather than behavior. Prevention models that target liability, like the ones described in this chapter, rather than specific behaviors are relevant to those engaging not only in the specific behavior targeted (i.e. drug use) but also in other addictive behaviors or maladaptive behaviors related to the personality profiles (i.e. such as gambling or antisocial behaviors in the case of impulsivity). Indeed a personality-targeted approach, targeting specific pathways to addiction, has already been shown to be effective in the prevention/early intervention of substance misuse, as well as other behavioral problems, in adolescence and adulthood. Although the effectiveness of this approach in the prevention or treatment of nonsubstance-related addictions has yet to be assessed, the consistent results obtained in the prevention of alcohol and drug misuse so far identifies this approach as promising and cost-effective, which can only improve current efforts in tackling substance use problems and other addictions so prevalent in society.

SEE ALSO

Behavioral Economic Factors in Addictive Processes, The Biopsychosocial Model of Addiction, Cognitive Factors in Addictive Processes, Emotions and Addictive Processes, Impulsivity, Disinhibition, and Risk Taking in Addiction, Interpersonal Factors and Addictive Disorders, Self-Medication, Adolescent Substance Use: Symptoms and Course, Symptoms and Course: Alcohol Use Disorder in Adulthood, Models of Relationships between Substance Use and Mental Disorders, Substance Use and Mood Disorders, Substance Use in Response to Anxiety Disorders

List of Abbreviations

5-HTT	serotonin transporter
5-HTTLPR	serotonin transporter gene-linked polymorphic region
AS	anxiety sensitivity
DRD2 and DRD4	dopamine receptors
H	hopelessness
IMP	impulsivity
MAOA	monoamine oxidase A
PET	positron emission tomography
SS	sensation seeking

Glossary

Affect Emotion or subjectively experienced feeling and its influence on behavior.

Anxiety sensitivity a personality trait characterized by fear of anxiety-related physical sensations due to an unrealistic expectation that they could lead to "catastrophic" consequences such as loss of physical or mental control.

Arousal a state of alertness and of high responsiveness to stimuli.

Comorbidity the simultaneous presence of two conditions or disorders in a patient or participant, e.g. substance use dependence and depression.

Cross-sectional a study design which provides information on the characteristics of, and statistical relationships between, variables at a specified moment in time.

Disinhibition the inability to restrain or prevent a behavior or impulse (opposed to inhibition).

Extraversion a personality factor characterized by interest in the outside world rather than the self, associated with traits such as sociability and assertiveness.

Factor analysis a statistical technique which examines the internal structure of a set of variables by analyzing the correlations between them, to identify underlying dimensions or factors.

Go/No-go task a task in which stimuli are presented in a continuous stream and participants are asked to either perform a motor response (go) or withhold a response (no-go) to specific stimuli.

Hopelessness a personality trait characterized by a tendency to have negative expectations concerning self and future life.

Impulsivity a personality trait often characterized by a deficit in reflectiveness and planning, rapid decision-making and action, and a failure to inhibit a behavior that is likely to result in negative consequences.

Inhibition the act or process of restraining or preventing a behavior or impulse, or the state of being "restrained."

Longitudinal a study design in which the same sample of research participants is examined repeatedly over time (over several time points).

Positron emission tomography medical imaging technique that produces three-dimensional images of the brain. This technique monitors regional cerebral blood flow in the brain by recording the emission of gamma rays when radioactively labeled glucose or some other substance introduced into the bloodstream is metabolized by neurons as they are activated.

Sensation seeking a personality trait characterized by a strong need for stimulation, a low tolerance to boredom, and a willingness to take risks for the sake of having novel and varied experiences.

Stop task similar to a go/no-go task, but one in which most stimuli require a motor response ("go" response), and a less frequent "stop" or "no-go" stimuli is presented immediately after the "go" signal, to which participants are asked to interrupt or inhibit the already-initiated motor response.

Further Reading

Blaszczynski, A., Nower, L., 2002. A pathway model of problem and pathological gambling. Addiction 97, 487–499.

Clark, D.B., Cornelius, J.R., Kirisci, L., Tarter, R.E., 2005. Childhood risk categories for adolescent substance involvement: a general liability typology. Drug of Alcohol and Dependence 77, 13–21.

Cloninger, C.R., Sigvardsson, S., Bohman, M., 1988. Childhood personality predicts alcohol abuse in young adults. Alcoholism: Clinical and Experimental Research 12, 494–505.

Conrod, P.J., Pihl, R.O., Stewart, S.H., Dongier, M., 2000. Validation of a system of classifying female substance abusers on the basis of personality and motivational risk factors for substance abuse. Psychology of Addictive Behaviors 14, 243–256.

Cooper, M.L., Agocha, V.B., Sheldon, M.S., 2000. A motivational perspective on risky behaviours: the role of personality and affect regulatory processes. Journal of Personality 68, 1059–1088.

Iacono, W.G., Malone, S.M., McGue, M., 2008. Behavioral disinhibition and the development of early-onset addiction: common and specific influences. Annual Review of Clinical Psychology 4, 325–348.

Leyton, M., Boileau, I., Benkelfat, C., et al., 2002. Amphetamine-induced increases in extracellular dopamine, drug wanting, and novelty seeking. Neuropsychopharmacology 27, 1027–1035.

O'Connor, R.M., Farrow, S., Colder, C.R., 2008. Clarifying the anxiety sensitivity and alcohol use relation: considering alcohol expectancies as moderators. Journal of Studies on Alcohol and Drugs 69, 765–772.

Pihl, R.O., Peterson, J.B., 1995. Alcoholism: the role of different motivational systems. Journal of Psychiatry and Neuroscience 20, 372–396.

Sher, K.J., Grekin, E.R., Williams, N.A., 2005. The development of alcohol use disorders. Annual Review of Clinical Psychology 1, 493–523.

Stewart, S.H., Devine, H., 2000. Relations between personality and drinking motives in young adults. Personality and Individual Differences 29, 495–511.

Tarter, R.E., Kirisci, L., Mezzich, A., et al., 2003. Neurobehavioral disinhibition in childhood predicts early age at onset of substance use disorder. American Journal of Psychiatry 160, 1078–1085.

Verdejo-Garcia, A., Lawrence, A.J., Clark, L., 2008. Impulsivity as a vulnerability marker for substance use disorders: review of findings from high-risk research, problem gamblers and genetic association studies. Neuroscience and Biobehavioral Reviews 32, 777–810.

Verheul, R., van den Brink, W., 2000. The role of personality pathology in the aetiology and treatment of substance use disorders. Current Opinion in Psychiatry 13, 163–169.

Zuckerman, M., 1999. Vulnerability to Psychopathology: A Biosocial Model. American Psychological Association, Washington, DC.

CHAPTER

29

Spirituality and Addiction

Thomas J. Johnson, Jean L. Kristeller

Indiana State University, Terre Haute, IN, USA

OUTLINE

BACKGROUND

Definitions of Religiousness and Spirituality

Defining religion or religiousness and spirituality (RS) has historically been a contentious issue, and it remains so today. Traditionally, functional definitions of religion focused on the why of religion; what does religion do for an individual or a group? Suggested functions have included supporting the social order, offering individuals a sense of meaning and purpose, prescribing codes of right living, or defending against the anxiety associated with our own mortality. Structural definitions focused more on what religion looks like; what are the necessary components or elements that make up religion? Such definitions often include some concept of ultimate goals, values, or truths and/or belief in the supernatural or in beings, powers, or principles that are more powerful, permanent, and/or transcendent than what humans typically encounter in the material realm.

Spirituality is being increasingly understood as a universal human capacity that involves experiencing a sense of meaning and purpose, a feeling of inner peace, feelings of awe, compassion toward others, and a connection with the sacred. Spiritual intelligence, although somewhat controversial as a construct, has been proposed as a way to conceptualize the pervasiveness of this capacity, the apparent range in regard to the degree individuals experience it, and the potential for cultivating such a capacity. One way that religion can be viewed is as culture-specific means to structure the needs engendered for understanding, codifying, and sharing the experiences labeled as spiritual.

Consistent with this, another contemporary approach is to equate spirituality with a search for the sacred (i.e. the ultimate, transcendent, or holy) and religion with

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00029-2

a search that occurs within some organized faith tradition. A related tendency is to equate spirituality with individual experience and religion with group traditions and rules. However, many religious traditions have been based on the spiritual experiences of their founders, and the stated goal of many faith traditions is to foster the spiritual development of adherents of that faith. Thus, RS are often seen as overlapping constructs rather than as opposites. Consistent with such a model, in national surveys in the United States, the majority of individuals report being both religious and spiritual, some see themselves as being religious but not spiritual, and a growing percentage report being spiritual but not religious. A few studies in the United States suggest that individuals in addiction treatment may be more likely than the general population to identify themselves as spiritual, but not religious, but broader samples are needed to determine how generalizable this finding is. Rates of church attendance and self-identification as being religious are generally lower in Europe than in the United States, but differences by country exist. In addition, in some European countries, spirituality is often conceptualized as related to Existential philosophy and questions of meaning in life.

Regardless of the relationship between religious belief and spiritual experience, it is important to recognize the variability that exists across individuals in the degree to which they identify with or experience these domains. Within the religious context, some people may identify themselves as religious but neither attend religious services nor espouse particular beliefs, other than general ones, such as in the existence of God; whereas for others, their religious beliefs and practices are core to their sense of self. Similarly, spiritual experience may vary from relatively common experiences such as the sense of well-being from walking through a natural setting or the feelings engendered while praying, to the intensity of a transformational or enlightenment experience that may occur only once in a lifetime.

In contemporary social science research, both spirituality and religion are commonly treated as multidimensional constructs. While global measures of one or the other can sometimes be useful, such research tends to focus on multiple aspects of RS. Key dimensions that have been identified and/or investigated in multiple studies include public religious behavior, private religious behaviors, importance of religious beliefs/faith to other aspects of self, spiritual or religious struggle, a sense of inner meaning and peace, the importance of values in shaping one's decisions, and the occurrence of mystical or transcendent experiences. Both spirituality and religion can each have both positive and negative relationships with substance use and abuse; furthermore, these relationships may differ depending on the type of substance involved (alcohol, smoking, other drugs of potential abuse, and compulsive overeating).

Psychoactive Substances and Religion

Psychoactive substances have been used in religious ceremonies for thousands of years. Sometimes substances were used to facilitate intense spiritual or transcendent experiences, such as use of hallucinogenic drugs in shamanistic ceremonies. Hallucinogens and other drugs were used to facilitate altered states of consciousness that the cultures engaging in these practices interpreted as entering or contacting the spirit world. Sometimes drugs were used along with other practices that were thought to facilitate such contact, including drumming, dancing, exposure to the elements, or abstaining from food, drink, or sleep for extended periods of time. Upon entering the spirit world, the Shaman could attempt to battle, banish, or appease spirits responsible for disease or other problems.

The use of substances to alter states of consciousness continues today. The Native American Church, incorporated in 1918 as a combination of native and Christian beliefs and practices, continues to use Peyote in highly structured rituals, but members generally do not appear to use the drug outside of the ceremonial context. Some Rastafarians use marijuana both as a part of religious ceremonies and outside such ceremonies. However, in both contexts marijuana is thought to facilitate spiritual wisdom and understanding of God. Several social movements in the second half of the twentieth century viewed drugs as part of a path to enlightenment and both inner and interpersonal peace. The psychedelic movement of the late 1960s advocated use of marijuana and lysergic acid diethylamide (LSD), while the Rave scene of the late 1980s and early 1990s combined use of 3,4-methylenedioxymethamphetamine (Ecstasy) with extended periods of dancing to loud, fast paced music. Sociologists do not agree on how to best classify or label such movements, but some have argued that they have quasi-spiritual elements.

Other religious uses of psychoactive substances have included giving substances as offerings to gods or spirits or assigning symbolic value to substances, such as association of wine with the blood of Christ in the Christian Eucharist. In contemporary settings, wine is used in religious rituals by Jews and some Christians. The amount of alcohol consumed in such rituals is usually not enough to have a noticeable effect on an individual. An exception is the Jewish holiday of Purim, where tradition dictates that one consumes alcohol to the point of confusion. Some faiths, especially Islam, some Buddhist traditions, and a number

of conservative Christian denominations (Latter Day Saints, some Baptists and Methodists, some Pentecostals) prohibit the use of alcohol as well as many other substances. Such prohibitions have been suggested as a way to mark religious boundaries or distinguish new religious movements from previously existing faiths. Faiths that prohibit alcohol use tend to have higher rates of abstinence than in the general population. However, in some, but not all, faiths that prohibit alcohol use, members who do drink may be at increased risk of alcohol abuse or dependence. Research in this area is complicated by the fact that not all members of such faiths are aware of or agree with their denomination's position on use of alcohol or other substances.

Religious and Spiritual Views on Addiction

In many Christian denominations, substance abuse problems have been labeled as sin. The nineteenth century approach to sinners of all types was often to try and persuade them to repent via threats of hellfire and damnation. While such approaches have not entirely disappeared, even in the nineteenth and early twentieth century, there were groups who took a more educational approach. The Emmanuel Movement started in 1905 in Boston, Massachusetts and attempted to use a combination of religious education and contemporary psychology in treating addiction and other problems. Today, Christian denominations vary widely in how they conceptualize sin. While some still attempt to inspire guilt to promote change, many others view sin as anything that comes between a person and God, and therefore focus their efforts on reconnecting a person to God. A related notion, common in various monotheistic faiths, is that addiction is a form of idolatry, placing something before God. Thus, the range of religiously based moral and theological positions on addiction is broad and complex.

Jewish writers note that alcoholism was once considered to be a problem that simply did not happen among Jews. While rates of alcohol dependence are typically lower among Jewish Americans than in the population at large, Jews can and do have problems with alcohol. Thus, a number of Jewish writers have worked to increase awareness of and provide supports for Jewish alcoholics. Authors from several other faith traditions have also expressed that when a religious group believes that drug or alcohol abuse is not likely to happen among their members, this can contribute to either a tendency to ignore such problems or to shun members who violate implicit or explicit norms, either of which could make it more difficult for such individuals to receive appropriate help for substance problems.

During the twentieth century, a common spiritual view of addiction has been that it is a spiritual quest gone wrong, stemming from individuals trying to find God in a bottle or seeking spiritual experiences via drugs. The prescribed solution then, is often giving the addict the correct religion or spirituality, which is, of course, defined differently by different groups. Much of the literature on RS and addiction has been in the form of personal narratives or case studies. However, even these reports suggest that there is no one single model or sequence for how RS is related to addiction or recovery. Some people may indeed begin using drugs as part of a spiritual search, but for others substance use might begin in a variety of ways, but become related to spirituality only when the substance problem becomes severe enough to interfere with a persons' spiritual or religious functioning.

In the United States, relatively few clergy from any denomination have had any training regarding addiction. In response to this the National Association for Children of Alcoholics (NACOA) and the National Institute on Alcohol Abuse and Alcoholism worked together to begin to create a training curriculum for clergy. NACOA has continued this work with some support from the Substance Abuse and Mental Health Services Administration, resulting in a variety of monographs and programs.

RESEARCH ON THE RELATIONSHIP BETWEEN RS AND SUBSTANCE USE

Numerous studies have identified inverse relationships between RS and substance use or substance use disorders (SUDs). Higher levels of religiousness, as measured by frequency of attending religious services or self-rated importance of religion, are generally associated with lower levels of substance use. However, prior to the 1990s, most such studies had used rather simplistic single item measures (e.g. frequency of religious service attendance, personal importance of religion, etc.). In contrast, contemporary research in the psychology of religion recognizes the importance of more sophisticated measurement of multiple aspects of RS.

The strength and even the direction of the relationship between RS and substance use can vary by culture, race, religious denomination, gender, and type of substance being used. For example, in the US personal religious commitment is generally inversely related to the amount of alcohol consumed or likelihood of being a cigarette smoker, but one national study found that in Lutherans, strength of personal religious commitment was actually positively related to heavy drinking. It is

not clear if findings such as this are related to aspects of religious belief systems, or if they could be due to other cultural traditions that might be correlated with religion (e.g. German family origin in American Lutherans, etc.). One specific aspect of religious beliefs that has been statistically linked to increased risk is the belief that God is harsh and punitive, as opposed to loving and forgiving. Individuals who hold a punishing view of God seem to show an increased risk for SUDs, but the nature of the causal relationship, if any, between these two variables is not yet clear.

Since the 1990s, more studies have attempted to identify mediators of the relationship between RS and substance use to clarify how RS might actually causally impact substance use. Beliefs about or attitudes toward substances, such as the belief that drinking alcohol is a sin, have some support as mediators, but the impact of beliefs may not be as strong as other factors, such as social influences or effects of RS on well-being. Some reviewers have suggested that RS beliefs may be more important in determining whether or not a person uses a particular substance than in affecting how much or how often a substance is taken among users. Furthermore, the impact of beliefs also appears to differ by religious denomination. In addition to the general finding that beliefs do not always predict behavior, there are a variety of reasons why beliefs about substances might not always show up as a strong predictor of the relationship between RS and substance use. In studies that collapse across denominations, the effect of beliefs for those in denominations that proscribe alcohol use might be diluted by the effects in denominations where such proscriptions do not exist. Further, averaged across all denominations, clergy in the US tend to believe that SUDs are a problem in their congregations, but seldom preach or offer information about substance use. This may contribute to the lack of knowledge of denominational policies regarding substance use evident in members of some denominations.

There is considerable evidence supporting the role of social influences as a mediator of the relationship between RS and substance use and SUDs. Social influences examined across studies have included modeling of use and nonuse behaviors, perceived norms for peer use, and access to substances. Both beliefs and social influences may affect substance use in part by reducing motivations to use substances as a way to have fun or socialize. Social influences also seem to have a direct effect on use, perhaps via reducing access to substances. It is also possible that some individuals self-select into particular peer groups based on their religious beliefs or patterns of substance use.

Another potential mediator of the relationship between RS and substance problems is well-being. Religious involvement tends to be associated with higher levels of well-being and positive affect, greater sense of purpose and meaning in life, and lower levels of negative affect. Thus, by enhancing well-being, RS may reduce the likelihood that individuals will use substances as a coping mechanism. There is in fact evidence from several studies that religious involvement reduces (moderates or buffers) the effects of stress on substance use. However, many of the studies that have focused on purpose or meaning in life used measures that are heavily confounded with depression, making it difficult to determine whether meaning in life is the key mediating variable, or some more general aspect of well-being. A number of other possible mediators have more limited support, including viewing the body as sacred and therefore not to be desecrated by unhealthy behaviors; helping others; and a sense of self-efficacy in resisting temptation.

Religious or spiritual struggles represent individuals' attempts to maintain or transform their connection to the sacred in the face of doubts, external stresses, negative life events, interpersonal conflicts within their faith community, or other problems. While doubts and struggles may be a common part of RS development, struggles seem to be particularly problematic when they persist and are not resolved over time. Substantial evidence connects struggles to negative affect and depression and a growing body of research indicates that struggles can also be related to SUDs. However, the direction of causality involved in the relationship between SUDs and struggles is not yet clear, and in fact may differ based on denomination or other factors. While some evidence supports religious and spiritual struggles as leading to the development of SUDs, in people coming from a denomination that proscribes alcohol use, using alcohol could itself be the cause of internal (e.g. feelings of guilt; self-condemnation) or interpersonal (e.g. criticism from co-believers, etc.) struggles.

RS AND MUTUAL HELP

Historically, religious organizations were among the first groups to attempt to provide treatment or support for individuals with SUDs. The Prohibition movement in the United States in the late nineteenth and early twentieth century was led by a vigorous coalition of primarily Protestant churches, growing out of concern related to the increasing social problems associated with virtually unregulated availability of alcohol. The development of mutual help groups, where addicts

help other addicts, has also been traced to religious movements as far back as the late eighteenth century.

Alcoholics Anonymous

The most well-known and widely available mutual help group is Alcoholics Anonymous (AA), founded by individuals who had been affiliated with the Oxford Group, an early twentieth century Evangelical Protest Christian movement. The founders of AA attempted to remove some of the more overtly religious elements of the Oxford Group, but aspects of that movement remain in AA practices such as the personal moral inventory, having a spiritual awakening, and spreading the movement to others. Twelve-step programs have also been developed for other substances (e.g. Narcotics Anonymous, etc.) and for behavioral addictions such as sex, gambling, and overeating. Most published research on mutual help has focused on AA. The dropout rate in AA is high, but exact rates are difficult to determine due in part to the organizations' emphasis on anonymity of members. Among individuals in treatment, degree of involvement in AA is a good predictor of positive outcome, but court mandated AA does not appear to be effective. Meta-analyses suggest a moderate effect size for the relationship between AA involvement and abstinence.

The role of spirituality in AA is widely cited, but empirical studies attempting to identify mediators of the effects of AA have not been able to document that having a spiritual awakening or other RS constructs explain the positive effects of AA. Rather, evidence supports a variety of constructs such as enhanced self-efficacy, increased motivation for abstinence, and finding a nondrinking peer group, as being the active ingredients by which AA helps some members. It is possible that the lack of evidence for RS mediators is due to existing studies doing a poor job of operationalizing and measuring RS in AA, so future work with more sophisticated measures would be helpful. Qualitative work identified an increased sense of spirituality and/or 'turning power over to a higher being' as an active component in the value of Overeater's Anonymous for individuals struggling with a pattern of compulsive or binge eating, but quantitative studies testing for statistical mediation are needed. Atheists and agnostics are less likely to affiliate with AA than individuals who have theistic beliefs. However, if atheists and agnostics do attend, they appear to derive as much benefit from the program as theists do. Some individuals have a negative response to AA because they view AA as being in conflict with their pre-existing religious beliefs. Some such individuals appear to affiliate with secular mutual help groups when those are available, while others may find help through mutual help groups specific to their faith.

RS and Other Mutual Help Groups

Following the spread of AA around the world, other mutual help groups were developed, often in reaction to the spiritual elements of AA. Secular groups such as Life Ring and SMART Recovery have attempted to remove spiritual elements from their programs. Perhaps paradoxically, some religiously committed individuals have been attracted to secular programs because they view AA spirituality as inconsistent with their personal religious faith. Other religious individuals have developed mutual help groups specific to their faith, including Catholic, Jewish, and Evangelical Protestant variations on the 12 steps. One of the most recent and widespread of these groups is Celebrate Recovery, a Christian recovery program founded in California in 1991. This program has spread worldwide, and is most popular in Evangelical and some Pentecostal churches. At present there is relatively little empirical research on either secular or faith specific mutual help programs. The Faces and Voices of recovery website maintains a list of secular and spiritual or religiously based recovery resources and mutual help groups, including the Buddhist Recovery Network; Jewish Alcoholics, Chemically Dependent Persons and Significant Others; and Celebrate Recovery.

RS IN THE TREATMENT OF SUDS

Meditation

In contrast to requiring abstinence or exerting self-control in the face of temptations, meditative or contemplative practices are intended to engage a higher level of awareness of the chain of addictive behaviors, thoughts, and feelings, as well as a gradual disengagement of the mind from desire for the addictive substance and related effects. Meditation practice can be grouped into three general types: concentrative, mindfulness, or guided. In concentrative practice, such as transcendental meditation (TM) or some Christian contemplative prayer practices, a word, sound, or mantra (often spiritually associated) or brief prayer is repeated to oneself. In mindfulness or Vipassana practices, a quality of quiet, detached awareness is cultivated through awareness of the breath or open, but nonreactive, awareness of whatever arises into consciousness. In guided practice, this quality of nonreactive awareness is directed toward a particular target, such as cravings, the quality of experience, or associated thoughts or feelings. Awareness of the breath, an aspect of many practices, is spiritually poignant; the word spiritual is derived from the Latin *spiritus,* which shares the root with *inspiration,* to breathe in. In Chinese (and Japanese) the word for spirit is *chi/qi* or *ki*, also the word for breath. In Jewish meditation,

incorporating the breath both metaphorically and as part of repetitive prayer is viewed as a way of taking in the spirit of God.

Even though contemporary meditation practices used for therapeutic purposes have largely been secularized, the traditional foundation for such practices was a spiritual/religious context in which it was understood that transformative processes were available to the serious practitioner. Furthermore, evidence suggests that even the secularized programs, such as Kabat-Zinn's mindfulness-based stress reduction program (MBSR), may increase a sense of spiritual well-being. Research over the last few decades has supported the value of a range of meditative practices for addressing and managing addictive behavior, and within the context of even relatively brief practice. A number of studies have investigated the effects of TM on both alcohol and drug dependency, including smoking. Use of alcohol and drugs appears to decrease, generally in relation to the amount of practice that is sustained.

Marlatt and his colleagues investigated two different formats to address drug and alcohol intake, within several different populations, including prison populations prior to release. Best known, and most easily translatable into general application, is his mindfulness-based relapse prevention (MBRP) program. The MBRP program melds his work on relapse prevention with principles of Buddhist psychology, including practices intended to cultivate a heightened sense of compassion and well-being. Although not explicitly engaging religious or spiritual themes, MBRP is intended to evoke spiritual engagement and cultivate these capacities. The other meditative context that Marlatt studied immerses individuals for a week within a silent traditional Vipassana retreat, developed by S.N. Goenka, a Vipassana teacher from Burma who has taught widely in the US and in India. This program has shown positive effects when used in a prison context, both in India and in the US. Notable in both programs is a flexibility in goals, such that abstinence is not considered the only optimal goal in regard to continued substance use. Rather, cultivating a great sense of control and wise choice is considered possible. This flexibility is also a hallmark of the mindfulness-based eating awareness training (MB-EAT) program developed originally for individuals with binge eating disorder. In contrast to Overeaters Anonymous, which maintains an abstinence model relative to certain types of foods, MB-EAT emphasizes cultivating awareness of "inner wisdom" to guide better food choices in regard to quality and quantity of food and the eating experience. Spirituality is engaged in this context, in regard to both the meditative experience and accessing a "higher" self to create better eating patterns. Evidence to date supports the value of MB-EAT in reducing binge eating and improving a sense of self-regulation.

Spiritual self-schema (3-S) therapy, while drawing on Buddhist principles and incorporating mindfulness meditation practice, is also tailored to each participant's own spiritual and religious beliefs. A variant of 3-S therapy, 3-S+, has been developed specifically for drug users who are HIV positive. When used in treating a group of HIV+ cocaine addicts, participants reported a greater sense of inner peace, increased spiritual faith, and a feeling of God's presence. In a randomized trial, impulsivity and addictive behaviors markedly decreased, and individuals tended to ascribe their improvement to both their mindfulness practice and to a deepening of their spirituality. Importantly, treatment manuals are available for both MBRP and the 3-S family of treatments, which should increase their dissemination and support further research on their efficacy with different populations.

Twelve-Step Facilitation Therapy

Twelve-step facilitation therapy (TSF) is a manualized therapy that focuses on helping the client become involved with a 12-step group and work the first three steps. In a large-scale national study of three manualized treatment approaches for alcoholism, TSF did at least as well as motivational enhancement therapy and cognitive behavioral therapy. TSF has also been explored in treatment of opiate and cocaine dependence, with at least preliminary evidence of benefit. While the first three steps addressed in TSF do include spiritual elements (believing in and surrendering to a higher power), there is as yet no strong evidence that religious or spiritual components of the intervention account for treatment effects.

Other RS Practices in Treatment

Prayer is often used in both twelve-step and faith-based (FB) (see below) programs, but relatively little research exists on how or if different types of prayer are related to treatment outcome or recovery. Complicating the picture is that there are a many different types of prayer. Petitionary prayer asks for God's help. Ritual prayer involves rote repetition of specific prayers. Intercessory prayer involves petitioning God to help some other person. One randomized trial of intercessory prayer found no difference between control participants and alcoholics who were being prayed for.

Some professionals working with individuals from a specific faith tradition have incorporated rituals or other elements specific to that tradition into the treatment process. For example, treatment programs for Native Americans may incorporate traditional symbols such as the medicine wheel, or traditional practices such as dancing and/or drumming. In some group

sessions an eagle feather is passed among the group members and whoever currently holds the feather is entitled to speak before the group. The published literature on spirituality in psychotherapy contains examples from other faith traditions, including Judaism and Christianity.

Given the interpersonal conflicts and complications that can occur in the lives of individuals struggling with addictive behaviors, several authors have suggested that forgiveness interventions may be helpful in addiction treatment. Early research in this area is promising, but more work is needed, because forgiveness, like prayer, is multifaceted. Forgiveness interventions can focus on forgiving someone else, forgiving one's-self, forgiving God, or asking for forgiveness by God.

Some studies have shown changes in various aspects of spirituality to occur over the course of addiction treatment, but more work is needed to establish that these changes mediate treatment effects. Studies have generally suggested that clients are open to discussion of spiritual issues in treatment, although this is liable to vary with the clients' specific faith tradition, specific type of addiction problem, or other client or cultural characteristics.

FAITH-BASED PROGRAMS

As noted above, religious groups have long offered treatment programs and various types of social services to individuals with substance abuse and other problems. For example, the Salvation Army worked with substance abusers from the very beginning of the organization. The House Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (HR 3734) included a provision for Charitable Choice, which allowed states to contract with both secular and sectarian FB groups to provide various types of social services. Since that time, Faith-based Organizations (FBOs) have received considerable attention from politicians, researchers, and the media. Focus on addiction services was heightened with 2003 access to recovery (ATR) program, which was touted as expanding consumer choice and increasing availability of addiction treatment and recovery support services (e.g. child care, employment services, case management, etc.). States receiving ATR funding were expected to contract with both FB and secular providers.

Supporters of FBOs noted that FB groups often work with populations that are marginalized and have limited access to conventional programs, sometimes due to geographical, cultural, or financial barriers. Criticisms of FBOs have included the concern that ATR and Charitable Choice would violate separation of church and state, as well as the concern that government regulations and paperwork requirements would actually interfere with FB programs and lessen their availability or effectiveness. Early proposed versions of Charitable Choice attempted to prevent states from requiring specific training for individuals providing addiction services. Such proposals received stinging criticism from the National Association of State Drug and Alcohol Abuse Counselors and the National Association of Alcohol and Drug Abuse Counselors, so that by 2006 there was general agreement that individuals providing addiction treatment in both secular and FB settings should be held to the same level of professional accountability. Perhaps the most overlooked issue regarding FB programs has been the question of how effective FB treatment and social services programs are. Supporters claimed that FB programs would be more effective than secular programs, but what little evidence is available regarding efficacy of FB social services does not support this contention.

Several factors cloud the picture on the efficacy of FB treatments for addiction. First, there is no commonly accepted definition of what type of treatment program counts as FB. Published studies on FB programs typically allow organizations to self-identify as FB. Secondly, organizations that self-identify as faith-based differ widely in terms of the degree to which religious elements influence their organizational structure and the services they offer, as well the degree of explicit and implicit religiousness present in the organization. In studies conducted over the first decade of the twenty-first century, relatively few self-identified FB programs relied exclusively on religious methods in treatment.

FB programs aimed at addiction treatment range from small, local ministries or missions to international organizations with multiple locations and programs. Perhaps the most widely known of the latter is Teen Challenge. Teen Challenge was founded in 1958 by Pentecostal minister David Wilkerson, who wanted to help Puerto Rican gang members he was ministering to in New York city. He initially assumed that a conversion experience would eliminate the problems youth were having with addiction and involvement in crime, but soon became convinced that a more intensive program was needed, as well as a more supportive environment. The original Teen Challenge model involved two phases. The first was several months at an Induction Center, followed by 6–9 months or more at a Training Center. The program involved bible study, religious education, and working, often in a rural farm setting. Today, the program largely houses and works with adults. No outcome studies of Teen Challenge have been published in peer reviewed journals, but the program has been open to outcome research. A federally funded study from the 1970s has often been cited as evidence for the efficacy of

Teen Challenge, but in fact dropouts from the program were as likely to be drug free at follow-up as were treatment completers. Given that the Teen Challenge program has changed considerably since its founding, outcome studies that are more than 30 years old probably cannot be generalized to the program today. A similar situation exists for Salvation Army programs, in that published outcome studies are 30 or more years old and of questionable methodological quality. Given the current emphasis on FB programs, the field is in dire need of more rigorous and recent outcome studies of FB treatments and social services.

CONCLUSIONS AND RECOMMENDATIONS

While religious involvement is generally inversely related to substance use and problems, even individuals who come from traditions that prohibit use of particular substances can develop SUDs. Religious or spiritual involvement may lower risk by contributing to beliefs that condemn use or reinforce moderate use, by increasing exposure to individuals who do not use and decreasing access to substances, and by enhancing well-being. Religious or spiritual struggles have been connected with substance abuse problems, but the direction of causation is not yet clear, as addiction could cause impairment in spiritual or religious functioning. While the inverse relationship between RS and addiction is supported by hundreds of studies, fewer studies have examined the role of RS in treatment. Meditation practices currently have the strongest support as helpful, but research continues on other RS approaches and practices. Three manualized approaches, TSF, 3-S Therapy, and MBRP, have preliminary support from more than one study and are worthy of more extensive investigation with a variety of populations.

Clinicians may wish to include spiritual assessment when planning treatment. The suggestions for further reading at the end of this chapter provide several sources that can assist in planning and carrying out such assessment. Even if religious or spiritual methods are not part of a treatment package, it may help to assess spiritual and religious functioning as part of a comprehensive approach to progress in and after treatment. Religious or FB organizations may also be sources of social services that may help facilitate recovery. Clinicians should become familiar with local FB resources, congregations, and clergy to more appropriately refer clients to programs and services.

Involvement in mutual help programs appears to be beneficial for many clients. Clients' preference for, or resistance to, specific mutual help groups may be related to the clients' own religious or spiritual beliefs. While 12-step groups remain the most widely available option, a growing number of secular and sectarian alternatives are also available and may be preferred by some clients. Clinicians should discuss with clients about their past experiences with mutual help, as well as how clients' specific faith tradition might relate to mutual help group participation.

SEE ALSO

Hallucinogens, Ecstasy/MDMA, Disease Model, Historical Understandings of Addiction

List of Abbreviations

AA	alcoholics anonymous
ATR	access to recovery
FB	faith based
FBOs	faith-based organizations
MB-EAT	mindfulness-based eating awareness training
MBRP	mindfulness-based relapse prevention
MBSR	mindfulness-based stress reduction
NACOA	national association for children of alcoholics
RS	religion and spirituality
3-S	spiritual self-schema
SUDs	substance use disorders
TM	transcendental meditation
TSF	twelve-step facilitation therapy

Glossary

Charitable choice a provision of the House Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (HR 3734), allowing states to contract with both secular and sectarian (faith based) groups to provide various types of social services.

Faith-Based programs religiously affiliated programs offering social services, health care, social support, or other types of aid.

Religion organized faith traditions that articulate principles of ultimate values and truths, provide belief systems that transcend the material realm, and promote experience of the spiritual.

Sacred holy, ultimate, transcendent, pertaining to a realm beyond that of our everyday lives.

Sacrament a ritual or ceremony that is given sacred meaning.

Spirituality a universal human capacity that involves feelings of inner peace, deep meaning, compassion, connection with the sacred, and related experiences.

Further Reading

Bowen, S., Chawla, N., Marlatt, G.A., 2011. Mindfulness-Based Relapse Prevention for Addictive Behaviors: A Clinician's Guide. Guilford, New York.

Carmody, J., Reed, G., Kristeller, J., Merriam, P., 2008. Mindfulness, spirituality, and health-related symptoms. Journal of Psychosomatic Research 69, 393–403.

Galanter, M., Kaskutis, L.A. (Eds.), 2008, Recent Developments in Alcoholism, Vol. 18. Springer, New York Research on Alcoholics Anonymous and Spirituality.

Gillum, R.F., 2005. Frequency of attendance at religious services and cigarette smoking in American women and men: the Third

National Health and Nutrition Examination Survey. Preventive Medicine 41, 607–613.
Johnson, T.J. Religion, spirituality, and substance use disorders. In: Pargament, K.I., Exline, J., Jones, J., Mahoney, A., Shafranske, E. (Eds.), APA Handbook of Psychology, Religion, and Spirituality, Vol. II. American Psychological Association, Washington, DC, in press.
Johnson, T.J., Bennett, P., 2009. Faith based programs. In: Cohen, L., Collins, F.L., Young, A.M., McChargue, D.E., Leffingwell, T.R. (Eds.), The Pharmacology and Treatment of Substance Abuse: An Evidence based Approach. Erlbaum, Mahwah, NJ, pp. 605–651.
Johnson, T.J., Sheets, V.L., Kristeller, J., 2008. Identifying mediators of the relationship between religiousness/spirituality and alcohol use. Journal of Studies on Alcohol and Drugs 69, 160–170.
Knipe, E., 1995. Culture, Society, and Drugs: The Social Science Approach to Drug Use. Waveland Press, Prospect Heights, IL.
Kristeller, J.L., Wolever, R.Q., 2011. Mindfulness-based eating awareness training for treating binge eating disorder: the conceptual foundation. Eating Disorders 19, 49–61.
Marlatt, G.A., Kristeller, J.L., 1999. Mindfulness and meditation. In: Miller, W.R. (Ed.), Integrating spirituality into treatment: Resources for practitioners. American Psychological Association, Washington, DC, pp. 67–84.
Mercadante, L., 1996. Victims & Sinners: Spiritual Roots of Addiction and Recovery. Westminster John Knox Press, Louisville, KY.
Michalak, L., Trocki, K., Bond, J., 2006. Religion and alcohol in the U.S. National Alcohol Survey: how important is religion for abstention and drinking? Drug and Alcohol Dependence 87, 268–280.
Nowinski, J., Baker, S., 2003. The Twelve Step Facilitation Handbook: A Systematic Approach to Recovery from Substance Dependence. Hazelden, Center City, MN.
Pargament, K.I., 2007. Spiritually Integrated Psychotherapy: Understanding and Addressing the Sacred. Guilford, New York.
Ringwald, C., 2002. The Soul of Recovery: Uncovering the Spiritual Dimension in the Treatment of Addictions. Oxford, New York.
Russell-Mayhew, S., von Ranson, K.M., Masson, P.C., 2010. How does overeaters anonymous help its members? A qualitative analysis. European Eating Disorders Review 18, 33–42.

Relevant Websites

For a list of secular and religiously affiliated mutual help programs:
www.facesandvoicesofrecovery.org/resources/support/resources/all.htm – Faces and voices of recovery.

Twelve-step groups

www.aa.org – Alcoholics Anonymous.
www.al-anon.alateen.org/ – Al-anon and Al-atten (for families of alcoholics).
www.ca.org – Cocaine Anonymous.
www.crystalmeth.org – Crystal Meth Anonymous.
www.doubletroubleinrecovery.org/ – Double trouble in recovery (for dually diagnosed individuals).
www.gamblersanonymous.org – Gamblers Anonymous.
www.heroin-anonymous.org – Heroin Anonymous.
www.marijuana-anonymous.org – Marijuana Anonymous.
www.na.org – Narcotics Anonymous.

Christian recovery and addiction related groups

www.addvicinc.org/ – Addictions victorius.
www.alcoholicsforchrist.com/ – Alcoholics for Christ.
www.alcoholicsvictorious.org/ – Alcoholics victorious.
www.calixsociety.org/ – Calix society (Catholic).
www.celebraterecovery.com/ – Celebrate recovery (Saddleback Church).
overcomersoutreach.org – Overcomers Outreach.

Addiction related groups from other faiths

www.jacsweb.org – Jewish alcoholics, chemically dependent persons, and significant others.
www.buddhistrecovery.org/links.htm – Buddhist recovery network.

Spiritual self-schema (3-S) therapy resources

info.med.yale.edu/psych/3s/index.html – The spiritual self-schema development program.

Other faith-related addiction resources

Faith partners – offers training programs for congregations of all faiths to help them better facilitate awareness of addiction problems and become more supportive communities for persons in recovery.
faith-partners.org/

The clergy education and training project – informational and training materials for clergy, developed largely by the National Association for Children of Alcoholics
www.nacoa.org/clergy.htm

So Help Me God: Substance Abuse, Religion, and Spirituality – A national study of clergy and faith groups conducted by the National Center on Addiction and Substance Use at Columbia University
www.casacolumbia.org/articlefiles/379-SoHelpMeGod.pdf

The Teen Challenge organization provides Christian-based residential treatment for addictions focused on teens and other age groups. teenchallengeusa.com/ – Teen Challenge.

CHAPTER

30

Interpersonal Factors and Addictive Disorders

Dorian Hunter-Reel
University of Washington, Seattle, WA, USA

OUTLINE

Interpersonal and social factors play an integral role in substance use, abuse, and dependence. Members of the individual's social network can be quite influential in their decision to initiate use of alcohol and other substances, continue substance use and/or abuse, seek treatment for problems, as well as to discontinue use. This chapter summarizes and explains the role of interpersonal and social factors in the initiation, maintenance, and resolution of substance use disorders.

INITIATION OF ADDICTIVE BEHAVIORS

While it is not universal, the majority of individuals who develop addictive behaviors initiate use during adolescence and early adulthood. However, the majority of research on the role of interpersonal processes on the development of addictive disorders focuses on adolescence. Research supports several theories about how social factors influence substance initiation.

Theoretical Models

Social Learning Theory

Social learning theory proposes that individuals learn by observing the behaviors of others (models). They then evaluate the effect of those behaviors by observing the positive and negative consequences that follow. Social learning theorists assert that members of the adolescent's social network who use substances serve

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00030-9

as models for adolescents. If adolescents see role models, such as parents or friends, using substances with positive consequences, they are more likely to develop positive expectations of substance use, which increases the likelihood that the adolescent will use substances. Learning not to use substances occurs in a similar fashion when adolescents who observe negative consequences of use expect negative outcomes and are less likely to use substances. Self-efficacy, an individual's confidence in their own capabilities, is also thought to be learned socially. Self-efficacy to obtain and use or, alternately, to refuse substance use, may also be learned by observing a model. If an individual perceives that he/she can efficaciously obtain and use substances, he/she is more likely to use them.

Social Control Theory

Social control theory holds that adolescents will engage in deviant behavior unless bonded to conventional societal institutions, such as family, schools, and religion, and to conventional role models such as parents and teachers. Adolescents with weak bonds to conventional institutions and role models fail to internalize the values held by conventional society. Such failure to form strong bonds may result from strain caused by poor relationships with parents or a discrepancy between an adolescent's goals and perceived ability to reach those goals. Social disorganization (the breakdown of established institutions that provide social control) may similarly result in failing to bond with conventional institutions and values. Adolescents without goals or family strain and with intact societal institutions available may not have been adequately socialized to adopt societal values, and may therefore turn to deviant peers, facilitating the development of substance use.

Social Developmental Model

The social developmental model is related to social control theory but emphasizes individual relationships. Theorists posit that adolescents are likely to become involved with substance using peers and use substances if they are disconnected from rewarding interpersonal and academic experiences. These adolescents may have failed to develop the interpersonal and academic skills that result in rewarding interactions with parents and teachers.

Theory of Planned Behavior

The theory of planned behavior asserts that adolescents choose to use substances based on both a cost–benefit analysis of the expected social and affective consequences, as well as their self-efficacy in using or refusing substances. Adolescents who have a strong desire to please peers are more likely to use substances if they perceive their peers using substances and want them to initiate use as well. They are less likely to use drugs and alcohol if they perceive risk and social disapproval resulting from substance use.

Contagion Theory

Contagion theory focuses on the spread of alcohol use among members of social groups. Research has shown that several health risk behaviors, such as smoking and drinking, spread contagiously throughout social networks. Changes in alcohol consumption by the members in an individual's network have a significant effect on subsequent alcohol consumption of the individual. These changes are not related merely to selective formation of ties with other drinkers. Changes in the consumption of an individual's relatives and friends predict similar changes in the individual's consumption at subsequent time points.

Network Member Influence

Peers

Adolescents often use alcohol and other drugs to create, maintain, and enhance social relationships. While adolescents who perceive their friends to be supportive of them are less likely to develop substance use disorders, adolescents who do not believe they have supportive friends are more likely to use substances as a means of garnering support or positive regard. Research examining motives for drug and alcohol use finds adolescents frequently report using alcohol and drugs because they believe using substances makes themselves and their friends feel better, alleviates interpersonal conflict between friends, and helps individuals to better fit in with the peer network. These kinds of beliefs tend to be particularly strong in adolescents with higher social anxiety (the fear of being embarrassed or humiliated in social settings).

Relationships also exist between substance abuse and certain kinds of social problems, including poor interpersonal skills, having few friends, and being a target of bullying. Adolescents with social problems are likely to have more positive attitudes toward substance use, greater difficulty refusing use, and a higher likelihood of using substances to escape social problems. However, other studies have found poor social functioning to be related to the absence of experimentation with alcohol and other substances. It is thought that adolescents who turn toward substances typically lack other means of coping. Adolescents with fewer cognitive distortions and adequate coping skills are less likely to turn to drugs and alcohol to cope. Adolescents with poor social functioning who associate with deviant friends are more likely to use substances versus those that do not. Associating with deviant peers increases the acceptance of

delinquent behavior and increases the availability of alcohol and other substances.

Parents

Adolescents who perceive their parents to be supportive and caring toward them are less likely to initiate substance use or develop substance abuse or dependence. This relationship is primarily due to the stress-buffering effect of effective parenting, support for academic competence, and lower tolerance of behavior problems. Parental warmth and affection, parental involvement and contact, parents use of healthy structure and discipline techniques, a relationship low in conflict, and adolescent attachment to and identification with a parent are all factors that have been shown to be associated with less use of alcohol or other drugs. Adolescents who perceive their parents to be tolerant of substance use are more likely to use substances, while those who believe their parents will punish them for using substances are less likely to initiate or maintain use.

The drinking and substance use behavior of parents is also important. In line with social learning theory, it has been found that very young children imitate drinking behavior they observe in their parents. Children of parents with substance use disorders are more likely to have problems with alcohol and other substances themselves. Alcohol and drug use disorders are thought to be largely genetically transmitted. However, having a substance-abusing parent has been found to confer risk for substance problems over and above the genetic risk. Children of substance abusers are even more likely to use substances if family rituals (such as eating dinner together, holiday celebrations) and social relationships were disrupted due to alcohol or drug use. The maintenance of family rituals is an important protective factor for children of addicted parents, particularly when the abstinence of the addicted individual is required for these rituals to continue. The relationship between parental and child substance abuse does not end once the child leaves the home. Adults who have continued relationships with parents with alcohol use disorders are more likely to have problems with alcohol themselves.

Siblings

Children and young adolescents spend more time with their siblings than they do with parents, friends, or even alone. Substance use by older siblings predicts initiation of substance use in younger siblings. Older siblings have a greater influence on substance use than either parents or friends. Several theories have been proposed to explain the similarity of substance abuse between siblings. Researchers have suggested that having an older sibling who uses substances increases the availability of drugs and alcohol to younger siblings, increasing the likelihood of early initiation. Younger siblings are more likely to select friends from their older sibling's deviant peer group, thereby increasing the availability of drugs and alcohol.

In line with social learning theory, researchers have asserted that adolescent substance use is transmitted by older siblings through the modeling of deviant behavior for younger siblings. Other researchers have asserted that the transmission of substance use behavior occurs by way of transmitting positive substance-related attitudes. This hypothesis is supported by the findings that older siblings are often reluctant to use substances in front of younger siblings, and in children who have not yet initiated use, the perceptions of their older siblings use and attitudes toward use are strongly associated with their attitudes and intentions to use. Perceptions and attitudes develop as early as 10 years old and predict later substance use. Therefore, exposure to sibling's substance-related attitudes and perceptions of their behavior may cause adolescents to have similar use patterns to their siblings, even if they do not directly observe substance use.

The "partners in crime" model holds that the closer the siblings are to each other, the more likely they are to influence each other in regard to using drugs and alcohol. Research has found that the degree of emotional warmth and support felt between the siblings, the amount of time they spend together, and the extent to which they have mutual friends all predict similarity in substance use.

As with parent–child relationships, it is important to note that at least some of the relationship between sibling substance use similarity is accounted for by genetic similarity. Studies of twins have shown that monozygotic (identical) twins have more social contact and mutual friends than dizygotic (fraternal) twins. However, dizygotic twins share social environment similarity to a greater extent than standard full siblings, demonstrating that some of the similarity in social networks is accounted for by genetic and some by age similarity. Shared environment also accounts for a substantial portion of social network similarity. However, when environmental factors such as family structure and income are accounted for, the similarity between sibling substance use remains significant.

Individual Characteristics

Certain social traits of the adolescents are associated with greater substance use. Adolescents high in social anxiety are more likely to develop alcohol use disorders. Motivation to cope and to conform to social expectations explains the association between social anxiety and

drinking problems. Moreover, adolescents with high levels of social anxiety also have greater consequences of alcohol use than non socially-anxious peers who consume the same amounts of alcohol.

Aggressive behaviors are associated with a variety of deviant behaviors in adolescence, including substance use. In adolescent males, this typically takes the form of physical aggression, which is a common predictor of substance use. For adolescent females, aggression typically takes the form of relational aggression, or deliberately causing harm to another by damaging their interpersonal relationships or social status. As with boys, relational aggression in girls is associated with substance use. It has been theorized that adolescents who have difficulty regulating their emotions and who have limited coping skills are more likely to engage in both relational aggression and substance abuse.

MAINTENANCE OF ADDICTIVE BEHAVIORS

Early Theoretical Models

Theories regarding the role of interpersonal factors and maintenance of addictions have developed over the last century. Models of family dysfunction in the maintenance of addiction first became prominent in the 1930s. The earliest theories were primarily psychodynamically focused. These theories attempted to explain depression, anxiety, and somatic complaints observed in the wives of men with alcohol use disorders by asserting that these women married alcoholic men to resolve their own neurotic conflicts. Similar theories asserted that these women depended on their husbands drinking as a defense mechanism and that they would decompensate if their husbands drinking problems remitted.

Beginning in the 1950s, research on the wives of men with alcohol problems contradicted these theories and suggested that the distress observed in these women was largely the result of living with an alcoholic and that wives and families developed strategies to cope with the stress of living with an alcoholic that are similar to patterns seen in families with a chronically ill member. It was also suggested that families go through stages in dealing with alcoholic family members, including denial of problem, attempts to control the problem, feeling hopeless and chaotic, attempting to maintain functioning, attempting to escape, organizing and maintaining the family without the alcoholic, and, if the alcoholic stopped drinking, a final readjustment phase. Further research has supported this theory – women with spouses with alcohol use disorders who are currently drinking are quite distressed in comparison to controls, while women married to men who have resolved their drinking problems have similar levels of distress as women married to men without alcohol use disorders.

Current Theoretical Models

Family Disease Model

The family disease model is an extension of the standard disease model, postulating that the entire family has a disease. While the disease model was born of the Alcoholics Anonymous (AA) movement of the late 1930s and early 1940s, the family disease model evolved with the establishment of Al-Anon in 1949 to support the family members of individuals with alcohol problems and became formalized largely in the 1980s. Several books were published asserting that codependence develops in children raised in households with an addicted parent, as well as spouses living with an addicted individual. Codependence is marked by "enabling" any behavior that perpetuates substance use or mitigates the effect of alcohol and drugs on the addicted individual. Typical enabling behaviors include making it easier to obtain alcohol or other substances, shielding the drinker from the negative consequences of their use, and keeping the addiction a secret. The family disease model has been widely influential and many modern treatments and popular conception are based on this model. However, relatively little empirical support has been found for the family disease model.

Family Systems Model

The family systems model of addiction originated in the 1970s and is similar to the family disease model in that it views the entire family unit as being affected. While the family disease model views enabling behavior as serving the function of maintaining the disease, the family systems model views the substance use as serving to maintain homeostasis. Under this theory, substance use is thought to be initiated and/or maintained during times when the affected family member is having difficulty with a developmental issue (for adolescents), or when the family is facing internal difficulties. The substance use then serves to distract from the core problem, or to inhibit developmental changes, thereby helping to maintain homeostasis within the family unit. The use of alcohol is thought to have an interpersonally useful role in facilitating communication, expression of affect, and intimacy within the family.

Behavioral Family Model

Behavioral family model is similar to the family systems model in its focus on functionality of substance use. However, the behavioral family model is based on learning theory and holds that substance use is learned

and maintained through principles of positive and negative reinforcement, which can include the reactions and behaviors of the family. Specifically, spouses and children may pay more attention, interact more positively, take care of a family member while he/she is under the influence, and alienate and punish the individual when they are sober, thereby reinforcing substance use.

Network Member Influence

A good deal is known about the social support networks of individuals with alcohol problems. Unfortunately, less research has focused on the social networks of individuals with other addictions. Therefore, relatively little is known about the networks of individuals with other substance use disorders. Within the alcohol-social support literature, a distinction is made between alcohol-specific and general social support. Alcohol-specific social support refers to support for abstinence and support for drinking whereas general social support refers to how generally supported an individual feels by their social network members. While both general and alcohol-specific social support are related to problem drinking, alcohol-specific social support has been found to be more strongly associated with it.

With regard to general social support, contact with social network members provides much of the basis for individuals' perception of support. The more social contact one has, the more supported they tend to feel. Individuals suffering from alcoholism are not typically without friends and support. Most have fairly large social networks averaging seven or eight close people. However, almost half of the individuals with alcohol use disorders live alone and/or far away from their families. This isolation is associated with heavier drinking, particularly for men.

With regard to alcohol-specific social support, about half of the individuals with alcohol use disorder have no one in their network that discourages drinking. Also, having more drinkers in the social network is associated with more frequent and problematic drinking. This is a particular problem for women with alcohol use disorders as they are more likely to have a partner with an alcohol use disorder. About half women with alcohol use disorders have a significant other with an alcohol use disorder, while only about 15% of men with alcohol use disorders do. While women do tend to have a lower proportion of drinkers in their social network than do men, they also tend to have larger social networks and see their drinking friends as more important to them. Men generally receive less support for drinking from their network members. Typically, only about a quarter of the social network members of men with alcohol use disorder encourage him to drink. For women, network members are more likely to encourage them not to drink, although women with moderate to heavy drinking husbands are more likely to receive encouragement to drink.

In line with the family systems and behavioral family models, some research has focused on the behaviors of the social network that may perpetuate drinking. Often this work has focused on how alcohol may restore homeostasis to the family. When there is tension, an individual with an alcohol use disorder may leave the home to drink, therefore relieving tension and perpetuating drinking behavior. The spouse may complain or argue about drinking when the individual is sober but not when he/she is drinking, a pattern which serves to reinforce drinking. Alcohol often serves as a tool to increase positive affect and leads to more positive experiences among family members. Steady drinkers and their spouses often show lower rates of psychological problems and higher marital satisfaction. However, this relationship has not been found with binge drinkers.

When drinkers have rewarding social experiences, they are more likely to drink in a social context to further facilitate positive social experiences. When drinkers have negative interpersonal experiences, they are more likely to drink alone, a pattern associated with the development of alcohol use disorders. This is particularly true for individuals high in neuroticism (those who are more likely to experience negative emotions).

Some research has suggested a feedback loop between social support and problem drinking. Less social contact is associated with lower levels of perceived social support. Lower perceived levels of social support are associated with high levels of depression. Higher levels of depression are associated with more drinking, and more drinking is associated with less social contact. Once this cycle has begun, it may be difficult for an individual with an alcohol use disorder to break out of it.

Individual Characteristics

While most theory and research has focused on how the social network influences individual's substance use, little work has focused on the interpersonal characteristics of the individual in relation to the social network in predicting substance use. Several interpersonal characteristics are known to be more prevalent among individuals with alcohol problems. Individuals with alcohol use disorders have poorer overall interpersonal functioning and have a harder time understanding the emotional facial expressions of others. Individuals with alcohol use disorders typically have more extreme scores on measures of problematic interpersonal behaviors, social withdrawal, and mistrust of others. Binge drinkers and those with severe alcohol problems are

more likely to engage in aggressive and violent behaviors toward their families and other network members. These interpersonal deficits and maladaptive behaviors may serve to further distance an individual from social support, perhaps further perpetuating substance abuse.

RESOLUTION OF ADDICTIVE BEHAVIORS

Interpersonal Predictors of Treatment Outcomes

Help-Seeking

Social network members frequently comment on the problem drinking of individuals with alcohol problems. Mothers, followed by spouses, friends, fathers, siblings, and children, most frequently make such comments. These comments are often beneficial. Individuals whose network members have spoken to them about their problem drinking are more likely to enter treatment. However, this pattern is somewhat less true for women, as they often receive less support or pressure from network members to seek treatment and are more likely to have a network member oppose their seeking treatment. While men often receive encouragement to seek treatment from spouses, when women receive such encouragement it is most likely to come from parents or children. Women are particularly likely to seek treatment if they have had interpersonal problems or difficulties fulfilling obligations to others as a result of drinking. Involvement with AA can also lead to greater formal help seeking, as individuals who attend AA meetings tend to receive more feedback about their drinking and encouragement to seek help and change their drinking.

Treatment Outcomes

Having supportive relationships with other people is associated with more positive treatment outcomes. Both general and alcohol-specific social support are important in predicting drinking outcomes of those being treated for alcoholism. Support for abstinence is associated with reduced drinking, and this relationship is magnified when general social support is also high. Similarly, when support for abstinence is low, the relationship between general social support and drinking outcomes is stronger.

Having a partner, being socially involved, having a larger social network, and having supportive network members, are all associated with better drinking outcomes for individuals being treated for alcohol use disorders. Both the degree of support from the most supportive person in the network and the number of people providing support strongly predict better drinking outcomes. Having even one supportive person in the social network is associated with less drinking during treatment.

With regard to alcohol-specific social support, a number of studies have found that support for not drinking predicts less risk of relapse and some have shown that support for drinking predicts a higher likelihood of relapse and a lower likelihood of becoming involved with AA. Having nondrinking social network members is associated with better drinking outcomes, and more support for abstinence from family, friends, and even work colleagues predicts a lower likelihood of relapse. Alternatively, support for drinking has not been as reliably linked to drinking outcomes as has support for not drinking. It has been shown that support for drinking from the most important people in the network is unrelated to drinking outcomes, while the overall amount of support for drinking from the entire network does predict drinking, perhaps suggesting that the individuals closest to the drinker are the least likely to support drinking.

The substance use status of social network members predicts outcomes. Having more drinking friends and maintaining relationships with them predicts greater likelihood of relapse. Having only one person in the social network who drinks or one network member who uses the same drug of abuse predicts greater likelihood of relapse.

Social Network Functioning

Individuals with alcohol use disorders who have good marital and family relationships, and social networks that provide support and assurance of worth are less likely to relapse. Individuals who relapse or fail to change their drinking typically receive more encouragement for drinking from their social networks. They also have more stressful friendships, experience more criticism and hostility from their families, view their families as being less supportive and cohesive, and have less marital satisfaction.

The spousal relationship is particularly influential. Negative interpersonal events involving the spouse are a strong proximal predictor of relapse. Passive responses to drinking such as withdrawing from the drinker, avoiding dealing with the topic of drinking, and tolerating drinking, are associated with poorer drinking outcomes of those being treated for alcohol use disorders. However, providing direct, specific, alcohol-related feedback is associated with less binge drinking among those being treated for alcohol use disorders.

Individual Differences

Social investment, the subjective value an individual places on his/her social network, is also an important factor. Social support for abstinence predicts better

drinking outcomes only for individuals who are highly invested in their social networks. Strong social support for abstinence is not important and may perhaps have some adverse effects for those who are not invested in their social networks. Age is also an important variable. While older individuals tend to have fewer drinkers in their social networks, they are more likely to change their drinking for the sake of relationships. The research on gender differences in the effect that social support has on drinking outcomes is mixed. Some studies have shown that women are more likely to be influenced by their social networks. Women are more likely to drink after interpersonal stress and conflict with their partners, and more likely to return to drinking with their partners than are men. Other studies have shown that men are more likely to drink after a stressful interpersonal interaction with their partners than are women.

As with maintenance of substance use disorders, little research or theory has focused on the role of the addicted individuals interpersonal behavior on the resolution of substance use problems. However, it has been found that individuals with domineering and vindictive interpersonal styles, and more overall interpersonal problems, are more likely to drop out of treatment for alcohol and other substance use disorders.

Changes in the Network Resulting from Treatment

After successful treatment, social networks often change. The ratio of drinkers to nondrinkers typically decreases and contact with friends increases. Women are particularly likely to have greater contact with friends and to view their friends as being important to them. For individuals who are involved with twelve-step groups, the social network is typically larger and composed of more abstainers. Those involved in twelve-step groups also tend to have more social network members who know about their drinking and substance problems, receive less social support for substance use, and view their friendships as being of higher quality.

Treatments Involving Social Influence

Treatments that directly involve family members have been shown to predict better compliance, treatment retention, and drinking outcomes. Involvement of significant others in treatment has increasingly been viewed as an essential part of substance abuse treatment.

Treatments Involving Social Network Members

A variety of family therapies have been developed to treat adolescent substance abuse. These treatments often focus on creating a developmentally normative lifestyle and helping parents to develop effective parenting styles. In comparison to individual and group treatments, family-based therapy has been shown to be more effective in reducing adolescent substance abuse.

Family systems therapy is based on family systems theory. It is designed to modify the family structure so that alcohol or other substance abuse no longer serves to maintain homeostasis. Family members are involved in therapy and encouraged to provide corrective feedback to the substance abuser and to create a family environment whose equilibrium is not based on substance abuse.

Behavioral family therapy is based on the family behavioral model and is similar to family systems therapy. Typically, the two basic goals of this treatment are to immediately eliminate and stabilize change in substance use and support the individuals efforts to change and then to alter family behaviors to reinforce continued abstinence. Individuals best suited for this type of treatment are typically older, with more severe dependence, who reside with an intact family whose members do not have substance-related problems or other mental health issues, and who have come to treatment as the result of a crisis threatening family relationships. Behavioral couple therapy is based on this same model, although it focuses specifically on the spousal relationship. In comparison to individual treatment, couple treatments lead to greater reductions in intimate partner violence, lead to better outcomes for children living with the couple, and is more cost-effective.

Some clients are more likely to benefit from this kind of approach than others. Individuals who have unsupportive social networks but are highly invested in them are most likely to benefit. For those uninvested inunsupportive networks, individual treatment is recommended. Conversely, if an individual is uninvested insupportive social network, they are likely to benefit more from the couple approach. If the patient is already invested in a supportive network, they are likely to receive maximum benefit from individual treatment.

An important note about family interventions: Clinicians conducting this type of treatment are trained to carefully clarify that it is the responsibility of the individual to change his/her substance use. However, spouses and family members can be helpful to the individual in his/her attempt to modify substance use, and all members of the family can learn how to interact with each other in more positive ways.

Interventions That Modify the Social Network

Mutual help groups have commonly been referred to as "self-help" groups. However, the term "mutual help" is more accurate as these groups are in large part organized around being helpful to each other with the common goal of abstinence, and individuals within the

group are encouraged to rely on each other for support. Researchers have also referred to these groups as "folk" treatments and "informal help" to distinguish them from formalized addictions treatment.

Twelve-step groups function in part by modifying the composition of the social network. Within the twelve-step philosophy, "carrying the message" and being helpful to other individuals suffering from addictions constitute an important element of the program. Twelve-step groups are anonymous and informal, and therefore difficult to study. Most of the extant research has focused on AA and has suggested that one of the primary reasons that it is effective is that it increases social support for abstinence, which, in turn, results in less drinking.

Several secular groups have since developed a focus on the ability of addicted individuals to be mutually helpful to one another without a spiritual focus. The Secular Organization for Sobriety (SOS) was founded in 1985. Women for Sobriety (WFS), an organization whose goal is to help women recover by sharing their experiences and encouraging each other, followed it.

Several programs utilize the group process, but use trained facilitators to lead groups and teach information or skills. Self-Management and Recovery Training (SMART) utilizes a facilitator to teach skills to cope with stress, rather than using substances. Moderation Management is similar to SMART, but emphasizes moderate use alcohol.

In comparison to twelve-step groups, there are far fewer secular and guided-help groups. Little research has been conducted on the efficacy of involvement in these groups, or how the interpersonal nature of the groups helps individuals change.

Treatments That Enhance Participation in Mutual Help Groups

Given the prevalence and efficacy of twelve-step groups, some treatments have focused on increasing involvement with mutual help groups. Twelve-step facilitation (TSF) typically takes the form of individual or group therapy explicitly designed to work synergistically with twelve-step groups. Therapists present addiction as a disease with spiritual, emotional, and physical aspects, consistent with twelve-step philosophy, and encourage involvement with twelve-step groups and "working the steps." TSF has been found to be as effective as cognitive behavior therapy and other top formalized treatments in achieving and maintaining abstinence.

Network support therapy is based on TSF, although it has an explicit focus on using AA to change the individual's social network to be more supportive of abstinence. AA philosophy is downplayed in favor of increasing AA attendance as a way to avoid drinking, make new friends, and engage in enjoyable activities other than drinking (to reinforce abstinence). The development of supportive social networks outside of AA is also explored. This treatment has been demonstrated to be effective in reducing drinking, as well as in increasing support for abstinence and increasing AA involvement.

Therapeutic communities use a social psychological approach to addictions treatment. Therapeutic communities are typically substance-free residential environments and utilize the "community as method" approach for addictions treatment. These communities offer a variety of rehabilitative services and focus on training members to adhere to social norms. Therapeutic communities rely on group members to give feedback to one another and to help each other develop more effective social skills. All of the members of therapeutic communities assume responsibility for their own recovery as well as the recovery of their peers. Individuals typically join these communities for 1–2 years and move up a hierarchy of increasing responsibility within the community. Research has shown that individuals who successfully complete these programs have significantly lower rates of substance abuse, criminal behavior, depression, and unemployment.

Treatments Initiated by or for Social Network Members

The Johnson intervention is an approach in which people close to the addicted individual work closely with a trained counselor to confront them, inform them of the negative effects of their use, and to ask them to enter treatment. While the majority of social networks who consider an intervention do not ultimately follow-through, among those that do, 75–80% are successful in convincing the individual to enter treatment.

Community reinforcement approach and family training (CRAFT), also known as unilateral family therapy, is a treatment for the social network member of the individual with a substance use disorder that teaches the network member to allow the addicted individual to experience the negative consequences of substance use and to reinforce abstinence. This treatment has been found to be effective for both moderating the distress of the social network member and encouraging the substance user to seek treatment. Over half of the individuals whose family members undergo this type of treatment later enter treatment themselves.

Al-Anon, Nar-Anon, Alateen, and Alatot are companion fellowships to AA for the social network members of individuals with substance use disorders. These groups parallel AA and use a modified version of the twelve steps, focusing on lack of control over loved ones' addictions, and emphasizing caring detachment. While involvement in these groups has not been shown to have an effect on the behavior of the addicted

individual, it has been shown to decrease the distress of the family member.

SUMMARY

Interpersonal factors and relationships play an important role in addictions. A variety of theories have been developed to explain the role that interpersonal processes play in the development, maintenance, and resolution of substance use disorders. Research has shown that social network members have a profound effect on one another. Behavior of network members may serve to protect individuals from the development of substance use disorders, or to increase the likelihood of developing and maintaining addictions. Social network members can also play an important role in promoting help seeking and aiding treatment efficacy.

SEE ALSO

Disease Model, Contextual Factors in Addiction, Families and Addiction, Peer Influences on Addiction, Epidemiology of Adolescent and Young Adult Alcohol, Tobacco, and Drug Use and Misuse in the United States, Drinking Patterns, Alcohol Consumption, and Aggressive Behavior

Glossary

Alcohol-specific social support the degree to which individuals perceive that social network members support either abstinence or drinking.

Behavioral family model based on learning theory, holds that substance use is learned and maintained through principles of positive and negative reinforcement, which can include the reactions and behaviors of the family.

Codependence a characteristic proposed to be present in family members of those with substance use disorders, marked by engaging in enabling behavior.

Contagion theory posits that health-related behaviors spread among members of social groups.

Enabling any behavior that perpetuates substance use or mitigates the effect of substance use on the addicted individual.

Family disease model an extension of the standard disease model, posits that the entire family unit has a disease.

Family systems model posits that substance use is serving to maintain relational homeostasis within families.

General social support the degree to which individuals perceive network members as being generally supportive.

Relational aggression the deliberate cause of harm to another by damaging interpersonal relationships or social status.

Social control theory holds that adolescents will engage in deviant behavior unless held in check by bonds to conventional societal institutions and role models.

Social developmental model posits that adolescents are likely to abuse substances if they lack rewarding interpersonal and academic experiences.

Social investment the subjective value an individual places on his or her social network.

Social learning theory posits that individuals learn by observing the behavior and apparent positive and negative consequences of social network members.

Theory of planned behavior posits that behavior is based on a cost–benefit analysis of the expected positive and negative consequences.

Further Reading

Hunter-Reel, D., McCrady, B.S., Hildebrandt, T., 2009. Emphasizing interpersonal factors: an interpersonal extension of an intra-individual model of relapse. Addiction 104, 1281–1290.

McCrady, B.S., 2004. To have but one true friend: implications for practice of research on alcohol use disorders and social networks. Psychology of Addictive Behaviors 18, 113–121.

Morris, E.P., Stewart, S.H., Ham, L.S., 2005. The relationship between social anxiety disorder and alcohol use disorders: a critical review. Clinical Psychology Review 25, 734–760.

Petraitis, J., Flay, B.R., Miller, T.Q., 1995. Reviewing theories of adolescent substance use: organizing pieces in the puzzle. Psychological Bulletin 117, 67–86.

CHAPTER

31

Families and Addiction

*Jeremiah A. Schumm**, *Timothy J. O'Farrell*[$]

[*]Cincinnati VA Medical Center and University of Cincinnati, Cincinnati, OH, USA,
[$]VA Boston Healthcare System and Harvard Medical School, Brockton, MA, USA

OUTLINE

INTRODUCTION

In order for clinicians and researchers to have a full understanding of the cause and treatment of addiction, it is important to consider the role of families in the development of and recovery from addictive behaviors. This article focuses on helping the reader to understand the association between major aspects of family functioning and their relationship with addiction. The article begins by describing the advantages and challenges in including family members into the process of assessing individuals with addictions, along with several major approaches to family-based assessment of addiction. We focus on family-based assessment and avoid discussion of specific family-based approaches to addiction treatment, since these are covered in other articles within this volume. The article then focuses on the association between addictive behaviors and relationship domains among romantic couples. We summarize major areas of research among romantic dyads, including couple's relationship satisfaction, behavioral communication patterns, and partner aggression and violence. Next, we turn to describing the empirical findings on the relationship between parental addictive behaviors and their children's adjustment. A summary of findings from the literature on adult children of alcoholic parents is also described. Finally, we provide suggestions for future research on the topic of families and addiction.

FAMILY-BASED ASSESSMENT

Incorporation of family members into the context of addiction-oriented assessment has several advantages. One advantage is that family members can serve as collateral sources of information as to the behaviors of the individual with the addiction. This can be helpful, as it is sometimes the case that the individual with the addiction will minimize his or her addiction behaviors and problems, due to stigma, shame, or other concerns.

Principles of Addiction, Volume 1
DOI: http://dx.doi.org/10.1016/B978-0-12-398336-7.00031-0

In addition, due to the difficulties that some classes of substances produce with regard to memory and cognitive functioning, asking a family member to provide collateral information as to the addicted individual's behaviors may help to increase accuracy of the information that is being collected. Another advantage of incorporating family members is that it can provide the opportunity to understand the role between family dynamics and addictive behaviors. As will be described throughout this article, there is compelling data to suggest that various domains of family functioning can influence addictive behaviors, and vice versa. Hence, assessment of family functioning and family member interactions may help to improve empirical and clinical models for understanding addiction. Lastly, incorporation of family members into assessment can provide the opportunity for their continuation in addiction-focused treatment. As is discussed in a companion article in this volume on Behavioral Couples Therapy for alcohol and drug use disorders as well as other articles in the volume, family-based interventions for addiction are shown to be effective, and there is some evidence that family-based interventions are more effective than individually based addiction treatment.

Although incorporating family members into addiction treatment and research has advantages, there are several challenges to obtaining information from family members. The first challenge is the willingness of the individual who is presenting with the addiction issue to consent to his or her family members being involved. Individuals with addiction issues may be fearful that inclusion of family members will lead to further embarrassment by family members revealing socially undesirable behaviors that occur within the context of the addiction. They may also be concerned that family members will bring up family-related problems that the person with the addiction would rather keep private. Another challenge is that the addictive behaviors may have contributed to family difficulties and resentment from the perspective of family members. Hence, family members may also be reluctant to want to involve themselves in providing information as it relates to talking about the addiction. Therefore, it is important for clinicians and researchers to provide clear expectations about what information will be communicated and how this information will be provided. To increase likelihood of family member involvement, it is also important to obtain consent for the researcher or clinician to speak with family members directly rather than relying on the individual with addiction to contact their family members. Further, we recommend explaining how involvement of family members will be helpful to them and their families and that the initial contact with family members be completed in the presence of the person with the addiction. A final challenge is that family members may be unable at times to serve as a reliable collateral source regarding the addiction behaviors. This may be due to individuals with addictions hiding these behaviors from loved ones. In this case, inclusion of the family member may not increase what is immediately known about the addictive behaviors.

There are three primary methods to consider when gathering information regarding family functioning. Some of these methods can be applied to gathering information on family functioning exclusively from the individual presenting with the addiction. However, as previously described, it is advantageous to collect data from multiple family members when using any of these methods. The first method to gather data on family functioning is to conduct a clinical interview. This method follows a practice that is widely used in general marital and family treatment. Clinical interview can be used to inquire about current family functioning, major family problems, and the role of the addiction in these, and the historical development of the addiction and family system over time. This approach has the advantage of allowing the assessor to explore in an unconfined manner nuances of the family system that may be relevant in understanding the addiction. This information can be useful, since relationship dynamics may play a role in the development of and recovery from addiction.

The second method for gathering information on family functioning involves the use of self-report questionnaires. There exist a multitude of widely used and well-validated questionnaires to assess various aspects of family functioning. A discussion of the advantages of specific questionnaires is beyond the scope of this article. However, use of questionnaires has a general advantage of allowing researchers and clinicians to obtain independent data from various family members in a standardized fashion. In addition, some family-based questionnaires provide normative-based comparison samples that allow for an interpretation of how individuals and families perceive themselves in relation to others on a given domain of functioning.

A third method that has both clinical and research application is to gather behavioral data from communication samples involving family members. Typically, this is done by asking family members to discuss a conflictual topic and recording observed behaviors and responses. There are methods for reliably coding these data (e.g. see work by Richard Heyman and colleagues). The usefulness of these data is that they provide microsamples of behavior with regard to how family members manage conflict among one another. As will be described later in this article and as described in the article on interpersonal factors in the addictive

process, methods for managing conflict are directly relevant to the development of and recovery from addiction.

ROMANTIC DYAD RELATIONSHIP SATISFACTION AND ADDICTION

One component of the family system that has been fairly well researched within the field of addictive behaviors is romantic dyad relationship adjustment. A common way of defining romantic dyad adjustment has been to measure overall levels of reported satisfaction within the couple relationship dyad. Relationship satisfaction has typically been construed as the degree to which partners feel supported by one another or are happy with the relationship. There are many measures of dyadic relationship adjustment, and some measures that have been commonly used in the addiction field include the Marital Adjustment Test and the Dyadic Adjustment Scale.

Numerous studies have examined relationship satisfaction among couples where one or both members exhibit an addiction. With regard to alcohol-related addiction, studies consistently show a negative association between having alcoholism and relationship satisfaction. For example, a study by Timothy O'Farrell and Gary Birchler compared 26 heterosexual couples in which the male partner had alcoholism to 26 nonalcoholic maritally distressed couples and 26 nondistressed couples from the community. They found that alcoholic couples reported less marital satisfaction versus nondistress couples and similar levels of marital satisfaction to nonalcoholic maritally distressed couples.

Consistent with the result that having a partner with alcoholism is related to lower relationship satisfaction, studies that have examined the frequency of heavy or problematic alcohol use have found a negative correlation between frequency of heavy or problematic drinking and couple relationship satisfaction. Further, according to a study by Mark Whisman, when compared to other mental health disorders, alcohol use disorder has one of the strongest negative associations with relationship satisfaction. Relative to the number of studies that examine the association between problematic alcohol use and relationship satisfaction, there has been less research on the relationship between illicit drug use and relationship satisfaction. Nonetheless, there is evidence that illicit drug use is also negatively associated with relationship satisfaction among couples. Finally, having a gambling addiction is also associated with less relationship satisfaction, although the number of empirical studies in this area is limited.

Although there seems to be an overall negative association between problematic substance use and relationship satisfaction, findings show that it is important to consider the pattern of substance use within the couple. Linda Roberts and Kenneth Leonard sought to identify how differing patterns of couple's drinking were associated with couple relationship satisfaction within a sample of newlywed heterosexual couples. Results showed that couples exhibiting heavy drinking on the part of the husband or heavy out-of-home drinking had lower relationship satisfaction versus those who reported lighter patterns of drinking, including those who drank together frequently but not heavily. Among a national sample of married and cohabitating couples, Katrin Leadly and colleagues found that couple discordance in partner drinking behaviors was associated with less relationship satisfaction versus when couples were concordant in drinking behavior.

In another large-scale longitudinal study of newlywed couples, Gregory Homish and others examined whether couples' drug use patterns were related to relationship satisfaction. They found the least female-reported marital satisfaction among couples where the partners had discrepant illicit drug use behaviors, followed by couples where both partners used drugs, and then couples where neither partner used drugs. These findings suggest that while drug use is associated with lower relationship satisfaction, discrepancy between partners with regard to drug use behaviors may have an especially detrimental effect on relationship satisfaction. Taken together, these findings suggest that specific patterns of substance misuse are likely to be associated with less relationship satisfaction.

ROMANTIC DYAD RELATIONSHIP COMMUNICATION BEHAVIORS AND ADDICTION

In addition to examining relationship satisfaction among couples with addictions, studies have also examined behaviorally coded measures of communication among couples. These studies generally show less constructive communication around conflictual topics for couples with addiction. Specifically, studies of couples in which one or both partners are alcoholic have generally shown higher levels of negative communication behaviors (e.g. hostility, criticism) and lower levels of positive communication behaviors (e.g. empathy, humor) among these couples versus nonalcoholic couples.

Theodore Jacob and Kenneth Leonard examined specific patterns of communication behaviors by comparing heterosexual couples with a male alcoholic partner to a demographically similar sample that was nonalcoholic. They found that in response to wives' problem-solving attempts during a communication sample, alcoholic husbands responded with less

positive and more negative behaviors in comparison to nonmaritally distressed, nonalcoholic husbands. In addition, when compared with nondistressed, nonalcoholic couples, wives in the alcoholic group were less likely to engage in problem solving and more likely to engage in negative behaviors in response to husbands' positive communication strategies. Findings among drug-abusing couples have also shown that when compared with nondrug-abusing couples, couples with a drug-abusing male partner tend to engage in negative communication styles that escalate over the course of conflictual discussions. Hence, when compared to those without addiction, couples with addiction exhibit a general pattern of higher dysfunctional and lower positive communication in response to conflict.

Similar to studies that examine overall relationship satisfaction as an outcome, there is evidence from alcoholic samples that contextual factors and partner characteristics may influence couples' communication styles. In an experimental study, Theodore Jacob and colleagues examined how drinking immediately before participating in a laboratory communication sample affected communication behaviors of heterosexual couples with a male alcoholic partner. This study also examined the degree to which husband antisociality affected couples' communication. They found that drinking before the communication sample increased negative couple communication behaviors among those with an antisocial husband. Wives of antisocial husbands were more likely to react with negative behaviors to their husbands' negative behaviors during the drinking versus no drinking condition. These findings suggest that having an antisocial male partner may increase the likelihood of couples engaging in negative conflictual communication during periods of drinking.

A more recent study by Frank Floyd and colleagues showed that partner discrepancy in alcoholism status may also be associated with couples' communication styles. The study compared couples on behaviorally coded communication tasks to see whether partner antisociality and partner alcoholism status influenced communication. They found that regardless of wives' alcoholism status, couples with an antisocial, alcoholic husband had the highest amount of hostile communication. Hence, similar to the earlier study by Jacob, male partner antisociality was an important variable in predicting negative communication.

Findings from Floyd and colleagues further showed that couples where the male partner was alcoholic but the female partner was not had the highest level of negative communication behaviors. Interestingly, couples where both partners were either both alcoholic or both nonalcoholic showed the highest ratio of positive-to-negative communication behaviors. This suggests that partner discordance in addiction status may moderate communication styles, such that those couples in which partners are concordant with regard to alcoholism status may exhibit better communication than those where the male partner is alcoholic but the female is not.

PARTNER AGGRESSION AND ADDICTION

Intimate partner aggression is shown to be fairly common among individuals with addiction. Studies suggest that the rates of past year intimate partner physical violence among those presenting for addiction treatment is between 50 and 60%. These rates are approximately five times higher than past year intimate partner violence prevalence among demographically similar, nonalcoholic individuals from the community. Intimate partner verbal aggression is also shown to be substantially increased among those entering addiction treatment versus demographically similar individuals from the community. Hence, intimate partner aggression is a significant problem among individuals with addictions.

The positive association between intimate partner violence and measures of alcohol and drug use behavior is well documented. Heather Foran and Daniel O'Leary conducted meta-analysis of 50 independent studies to examine the association between intimate partner violence and alcohol use. Findings from this study showed a small, but statistically, significant association between female-to-male intimate partner violence perpetration and female alcohol use. The overall effect size for male-to-female intimate partner violence perpetration and male alcohol use was also significant and was in the small-to-medium effect size range. These results were further qualified by the alcohol–aggression association being dependent on the type of alcohol use measure that was used. Namely, measures of problem drinking produced a larger effect size with aggression in comparison with measures of consumption. Also, measures of drinking quantity produced higher effect size associations with aggression versus measures of drinking frequency. Not surprisingly, these results indicate that more problematic drinking patterns have a stronger association with partner violence.

A meta-analysis by Todd Moore and colleagues examined the relationship between drug use behaviors and intimate partner aggression. Findings from 96 studies indicated a small-to-medium effect size association between drug use behaviors and intimate partner aggression. However, differences were noted in the degree of association, depending on what type of aggression was being considered. Results indicated a significantly lower effect size association between

drug use and psychological aggression versus drug use with physical, sexual, or mixed forms of aggression. Regarding the type of drug that was being assessed, small effect sizes were noted for most drug types, whereas cocaine had a small-to-medium effect size association with aggression. The degree of association between cocaine and aggression was significantly higher versus several other types of drugs and aggression. The pattern of results suggests that while various classes of drugs seem to increase the risk for partner aggression, cocaine may have a more substantial relationship with partner aggressive behaviors relative to other drug classes. Although gender of the drug user was examined as a moderator, no differences were found in the effect size relationships between drug use and partner aggression.

Findings from large-scale, prospective community-based studies suggest that substance use increases risk for intimate partner aggression. In a 3-year, longitudinal study, Kenneth Leonard and Marilyn Senchak found that marital conflict styles along with husband alcohol use were prospectively predictive of later husband violence. In addition, a large-scale, longitudinal study by Maria Testa and others showed that women's drug use increased the likelihood of their experiencing partner violence in both ongoing and new romantic relationships.

Studies have also been conducted to examine variables that differentiate partner violence from nonviolent couples who are diagnosed with an addiction. Both proximal (e.g. intoxication) and distal (e.g. personality characteristics) variables may influence the occurrence of partner aggression among couples with an addiction. In a study of proximal risk factors for partner aggression, Christopher Murphy and colleagues interviewed alcoholic men and their female intimate partners to determine what circumstances were associated with conflicts that escalated to aggression. Participants endorsed the husband's alcohol use as the main topic in the majority of conflictual occurrences. With most conflicts, husbands were reported to have been consuming alcohol during the 12 h preceding the conflict. Hence, it seems that conflicts are typically related to the problematic drinking behavior, and conflicts are likely to occur during drinking episodes. In addition, when violent conflicts arose, husbands were reported to have consumed a higher quantity of alcohol. This may be related to the manner in which higher levels of intoxication lower inhibitions toward socially undesirable behaviors, such as violence.

Other studies have examined distal risk factors for partner violence, including individual and couple-based characteristics among those seeking treatment for alcohol or drug addiction. Studies have generally shown that higher levels of alcohol and drug use and higher levels of substance-related problems increase risk for partner aggression. In addition, an individual characteristic that has been researched among men but not women with addiction is antisociality. Among alcoholic men, higher levels of antisociality are associated with increased likelihood of partner violence perpetration. With regard to relationship factors among couples with addiction, partner violent couples are shown to exhibit less relationship satisfaction and a stronger perception that the addiction is causing relationship problems. In addition, alcoholic couples who engage in partner violence are shown to exhibit poorer behaviorally coded communication strategies versus those who are not violent. In summary, among individuals seeking addiction treatment, having higher addiction problem severity, having an antisocial male partner, and having lower relationship adjustment seems to distinguish couples who are violent from those who are nonviolent.

In a national community sample of couples, Katrin Leadly and others examined whether couple drinking patterns were associated with partner violence. Their results showed that couples exhibiting moderate or frequent drinking had over twice the odds of experiencing partner violence versus couples who abstained from drinking. In addition, couples who exhibited discrepant drinking patterns had over three times the odds of experiencing partner violence than those who abstained. These results were consistent with the pattern of findings from clinical samples that suggest that partner discordance in addiction-related behaviors is related to a range of couple-related problems, including partner violence.

PARTNER'S INFLUENCE ON ONE ANOTHER'S ALCOHOL USE

Understanding the longitudinal course of addictive behaviors within intimate dyads is important because it can increase the understanding of factors that influence the development of addiction, increase the ability to identify those at high risk, and improve the development and refinement of addiction interventions. In a longitudinal study, Kenneth Leonard and Pamela Mudar followed couples shortly after becoming married and through their first year of marriage. In this study, peer and partner drinking, along with other known risk factors for alcohol problems, were examined to determine whether these influenced husband and wife alcohol use behaviors during their first year of marriage. Findings indicated that heavier husband drinking when the couple initially became married was predictive of increased wife's drinking behaviors at their first year anniversary. However, wife's drinking behaviors when

the couple was initially married had no effect on husband's drinking behaviors at their first year anniversary. In addition, peer drinking behaviors when initially married did not affect either partner's drinking behaviors at their first year anniversary. In contrast, heavier husband drinking when the couple was first married was predictive of increased husband and wife involvement with alcohol-using peers at their first year anniversary. These results suggest that husbands' drinking habits seem to drive the drinking habits of their wives early in marriage and influence the degree to which couples engage with an alcohol-using peer group. These findings suggest that targeting husband drinking behaviors early in marriage may be a key component to addressing both partners' alcohol use and potential abuse.

PARTNER INFLUENCE IN THE RECOVERY FROM ADDICTION

Studies suggest that partner behaviors can affect the chances of recovery from addiction. Building on the literature on expressed emotion (EE) and its relationship to relapse in nonaddictive mental health disorders, Timothy O'Farrell and colleagues examined the role of female partner EE in predicting men's recovery from alcoholism after treatment. After controlling for other known relapse risk factors, they found that when compared with those who had low EE partners, alcoholic men who had partners who were high in EE were more likely to relapse, relapsed quicker, and had more heavy drinking during the year after treatment. Similar results were also found in a more recent study. In this more recent study of male alcoholic individuals, Mattson and colleagues found that female-perpetrated psychological aggression toward men with substance use disorders increased the likelihood of relapse during the 6 months after treatment. This result was found after accounting for the negative association between overall relationship satisfaction and relapse, as well as demographic and substance-related risk variables. This suggests that the relationship between EE and psychological aggression with addiction recovery are not a simple by-product of overall relationship adjustment or other recovery risk factors. Similar results have also been found in drug-addicted men, such that higher perceived criticism from female partners predicts higher likelihood of men relapsing to drugs. Hence, partner criticism and psychological aggression seem to decrease the chances that individuals with addiction will be able to sustain sobriety after treatment.

In addition to partner communication strategies influencing likelihood of recovery, partners may unintentionally engage in behaviors that enable the continuation of addictive behaviors. In a sample of heterosexual couples in which one partner exhibited addiction and the other did not, Rob Rotunda and others found that the vast majority of nonaddicted partners engaged in one or more enabling behaviors at some point during the relationship. Some examples of common enabling behaviors included taking over of chores for the addicted partner, drinking or using drugs with the addicted partner, or making excuses to cover the addicted partner's drinking or drug use behaviors. Unfortunately, such behaviors may unintentionally promote the continuation of the addiction by reinforcing addictive behaviors. Hence, these results suggest a clear need to address family member enabling to increase chances at addiction recovery.

THE RELATIONSHIP BETWEEN PARENTAL ADDICTION BEHAVIORS AND THEIR CHILDREN'S OUTCOMES

In addition to examining the association between addictive behaviors and romantic partner relationships, there has also been research to investigate the association between parental addiction and behaviors of their offspring. In this section, we focus on studies that examine how parents' addictive behaviors may affect their offspring's behavior during childhood. However, there is also a body of research examining the intergenerational transmission of addiction and the long-term effect of addiction across multiple generations. For discussion of the intergenerational aspect of addiction and its consequences, we refer interested readers to the article by Campbell in this volume.

There is clear evidence that children of individuals with addictive disorders who are at increased risk to themselves exhibit childhood behavioral disorders. In a large-scale, community sample of adolescent twins, Marmorstein and colleagues found significantly higher rates of various childhood disorders, including conduct disorder, oppositional deviant disorder, attention deficit disorder, and substance use disorders among adolescents who had a parent with alcohol dependence versus those whose parents did not exhibit alcohol dependence. Similarly, having a parent who was dependent on a drug other than cannabis was associated with increased odds of having these childhood disorders as well. Regarding prevalence of specific childhood disorders, the highest prevalence rates were found for oppositional deviant disorder, conduct disorder, and nicotine dependence, such that approximately one-third of the children of a parent with drug dependence other than cannabis exhibited these disorders. Prevalence for these disorders among children of parents with noncannabis drug dependence was almost 10 times the prevalence found

in children whose parents were not dependent on alcohol or drugs. Among children with an alcohol-dependent parent, it was most common to exhibit conduct disorder, with one-quarter of these children having a conduct disorder.

Other longitudinal research supports a connection between alcoholic fathers' drinking and their children's behavioral problems. Andrea Burdzovic Andreas and Timothy O'Farrell examined the frequency of alcoholic fathers' heavy drinking and child behavioral problems in the 12 months before fathers engaging in alcoholism treatment, during treatment, and in the 12 months after treatment. Children in this study ranged from ages 5 to 16. Cluster analysis revealed three patterns of alcoholic fathers' heavy drinking frequency during the course of the study: stable and low, low and increasing, and high and increasing. Children of fathers whose drinking was classified as either stable and low or low and increasing showed significant declines in behavioral problems from pretreatment through the 6 months after treatment. It was noteworthy that children with fathers who exhibited a low and increasing drinking pattern had the most reduction in behavioral problems during their fathers' periods of abstinence. In contrast, children of fathers whose drinking was classified as heavy and increasing showed no changes in behavioral problems from pretreatment through 6 months after treatment, but then had increased behavioral problems from 6 to 12 months after treatment. These results suggest that alcoholic fathers' abstinence is positively related to their children's adjustment over time, whereas continued heavy drinking is associated with increased risk for child behavioral problems.

Other research by Burdzovic Andreas and colleagues demonstrates that fathers' involvement in addiction treatment and self-help may lead to reductions in their children's behavior problems. Children in the study sample had an average age of 13, and the vast majority were age 13 or younger. Results showed that the extent of alcoholism treatment participation predicted posttreatment self-help involvement. Higher self-help involvement after treatment was, in turn, related to fathers achieving higher abstinence after treatment. Alcoholic fathers' abstinence was associated with their children having fewer externalizing behavioral problems after treatment. These researchers also compared children of alcoholic fathers who participated in alcoholism treatment with a demographically matched community sample. Their findings showed that children of alcoholic fathers who remained stably remitted during the 15 months after treatment had significant reductions in externalizing behavioral problems after treatment. In addition, the degree of behavioral problems exhibited by children of remitted alcoholic fathers did not differ from those found in the community sample and were lower than those whose alcoholic fathers had relapsed after treatment. This suggests that addiction treatment and self-help involvement may have collateral benefits for younger, nonadolescent children of individuals with addictions.

Among adolescent children of parents with addictions, there is evidence that addiction remission might mitigate, although not entirely eliminate, the risk for behavioral problems. Laurie Chassin and her research team compared adolescents who had a parent with an alcoholism history with a demographically similar group of adolescents whose parents were nonalcoholic. They found that having a history of parental alcoholism increased risk for adolescents exhibiting internalizing and externalizing behavioral problems and substance-using behaviors. This risk was moderated by the recency of parental alcoholism, such that adolescents whose parent exhibited alcoholism problems within the past 3 years had higher relative risk for internalizing and externalizing behavioral problems versus those whose parents did not have an alcoholism history. In addition, those whose parent exhibited alcoholism problems within the past 3 years had higher relative risk for substance use and problems related to their use versus those whose parents were nonalcoholic. In contrast, those whose parents had a remote history of alcoholism problems before the past 3 years did not differ on risk for externalizing behavioral problems versus those whose parents were nonalcoholic. Adolescents who had a parent with a remote alcoholism history continued to exhibit higher risk for internalizing problems, based on their fathers' reports and had higher risk for recent substance use and lifetime substance-related problems. Hence, the risk for externalizing problems was reduced among those whose parents had remote alcoholism history, while an increased risk for internalizing problems and substance use was present for adolescents with both a recent and remote parental alcoholism history.

In addition to literature supporting a relationship between parental addiction and child behavioral problems, there are also studies suggesting that parental addiction increases the risk for their children being exposed to violence. Research shows that children of parents with addictions have higher rates of exposure to intrafamilial violence, as well as higher rates of exposure to violence that occurs outside of the home. Steven Ondersma and colleagues examined this issue specifically among African American women and their children. They found that after accounting for multiple other factors, including children's behavioral problems, mothers' stress and depression, and neighborhood crime, exposure to alcohol and drugs was

related to children's higher violence exposure. Hence, these results suggest that exposure to alcohol and drugs may confer a unique risk to children having experiences with violence. Although there are known detrimental consequences for children with parents who have addictions and those who are exposed to violence, the combined effects of these risk factors is not well studied. This is unfortunate given that it is well established that addictions and violence commonly co-occur, and these risk factors occurring together are likely to have a detrimental effect on child development.

PATHWAYS FROM PARENTAL ADDICTIVE BEHAVIORS TO THEIR CHILDREN'S BEHAVIORAL PROBLEMS

The relationship between parental addictive behaviors and their children's behavioral problems may be explained by several pathways. In a 3-year longitudinal study, Laurie Chassin and others investigated familial factors that predicted changes in adolescent substance use over time. Using latent growth modeling, they found that having an alcoholic biological father predicted increases in adolescent substance use during this 3-year period. In addition to this direct effect of father alcoholism status, these researchers found that mother and father alcoholism status indirectly influenced adolescent substance use through several complex pathways. For example, having an alcoholic mother was related to less father monitoring behaviors. Less father monitoring behaviors was associated with higher involvement with drug use peers, and involvement in drug-using peers, in turn, predicted increased substance-using behaviors over time. Results suggested that having an alcoholic father also indirectly increased adolescent substance-using behaviors through this and other indirect pathways. Hence, parental alcoholism status might indirectly affect adolescent substance use behaviors through a chain of pathways that lead adolescents to engage with substance-using peers, thereby increasing the likelihood of substance use.

In addition to findings that support the role of parent–child relationships in explaining the link between parental alcoholism and their adolescent children's behaviors, there is evidence from research on younger children that parent–child relationships are important in explaining the effect of parental problem drinking on children's behavioral problems. In a sample of children aged 6–12, Mona El-Sheikh and Joseph Buckhalt found worse family cohesion and adaptability and worse child behavioral problems among alcoholic versus nonalcoholic families. However, children's perceptions of attachment to their parents moderated the link between parental alcoholism and childhood behavioral problems. Specifically, securely attached children were only at increased risk for problems when they had a parent with alcoholism and did not exhibit increased risk for problems when neither parent was alcoholic. This suggests that parental alcoholism may be an important determinant of child problems among securely attached children, whereas insecurely attached children may be prone to exhibit problems even if their parents are not alcoholic. In addition, family adaptability and cohesion buffered the effect of parental alcoholism status on their children's problems such that children who were in families with higher adaptability and cohesion were less negatively affected by having a parent who was alcoholic. These findings show that if families are able to maintain a well-functioning family system despite the challenges introduced by parental alcoholism, they may be able to partially offset the detrimental effects of parental alcoholism on their children's behaviors.

ADULT CHILDREN OF ALCOHOLIC PARENTS

Although the effect of being an adult child of alcoholic parents has received significant attention in the popular media and among some clinicians, drawing conclusions from the research literature is difficult, since studies of adult children of alcoholic parents have mostly relied on adult children of alcoholic parents providing retrospective reports of their parents' drinking behaviors and retrospective reports of their childhood environments. Hence, the causal relationship between childhood factors, including parental drinking during childhood, and adult children's adjustment is unclear.

Nonetheless, studies that have been conducted on adult children of alcoholic parents have tended to show that these individuals have increased risk maladjustment in a number of domains. According to Stephanie Harter's review of the literature, adult children of alcoholic parents are most consistently shown to exhibit increased risk for substance use disorders, antisocial behavior, emotional disorders, low self-esteem, and family problems. Although they seem to be at increased risk for maladjustment, Harter concludes that there is no evidence to support an "adult children of alcoholics" syndrome. Rather, adult children of alcoholic parents seem to be a heterogeneous group that exhibit increased risk for various forms of maladjustment. Although having an alcoholic parent is related to increased risk for adulthood problems, it is not clear the degree to which this increased

risk in having an alcoholic parent is distinct from having a parent with another mental disorder or being exposed to a generally dysfunctional environment in one's family of origin.

In a rare, large-scale, longitudinal study on adult children of alcoholic parents, Karen Jennison and Kenneth Johnson investigated whether having an alcoholic parent and family environmental factors predicted young women's development of alcohol use disorders. They followed women over a 10-year period and found that having an alcoholic parent did increase risk for adult children to develop an alcohol use disorder. However, the magnitude of having an alcoholic parent on their adult children's alcohol use disorder markedly decreased as the children grew older into their twenties and thirties. In addition, the effects of women's marital relationships moderated the influence of having an alcoholic parent on women's alcoholism outcomes. Specifically, women who reported more cohesion and better communication in their current marital relationships were less affected by having an alcoholic parent. Hence, having a supportive and well-functioning marital relationship seems to be an important resource for buffering the risk of having an alcoholic parent.

SUMMARY AND AREAS FOR FUTURE RESEARCH

In summary, there is a clear negative association between addictive behaviors and family functioning. Addictive behaviors are associated with a variety of romantic partner relational problems. Specifically, when compared to couples where neither partner has an addiction, couples in which one or both partners exhibit an addiction are shown to exhibit less relationship satisfaction, poorer communication, and more intimate partner aggression. However, studies also show that it is important to consider the pattern of addictive behaviors between partners, as those couples where one member exhibits addiction and the other partner does not are shown to have worse dyadic functioning versus those where both or neither partner has an addiction. In addition, there is evidence that partner behaviors and relationship factors may help or inhibit the likelihood of recovery from addiction. These findings suggest that partners' roles are clearly important to understanding the course of and recovery from addiction.

Regarding the effect of parental addiction on children, parents' addictive behaviors are shown to be positively related to child behavioral problems. Parents with addictions have children who exhibit a higher risk for mental health disorders versus parents who do not exhibit an addiction. Further, parental addictive behaviors are longitudinally related to childhood behavioral problems, such that more frequent problematic substance use predicts more childhood behavioral problems. Likewise, recovery from addiction and reduction of parental addictive behaviors is related to fewer childhood behavioral problems over time. Studies suggest that dysfunctional parenting behaviors and parent–child relationships may be one pathway through which parental addiction affects children.

Although the connection between addictive behaviors and the family is well established, there is still need for expanded research among various populations exhibiting addiction and its relationship to family dynamics. Family-based addiction research has historically focused on heterosexual dyads where the male partner exhibits an addiction. Additional research is needed to expand the understanding between addiction and family functioning among a more diverse range of populations, including women exhibiting addiction and homosexual couples. Additional research is also needed to examine how family functioning and addictive behaviors unfold over the course of time. This would improve the understanding of the degree to which addictive behaviors erode family functioning and how positive family functioning may serve as a resource to promote recovery.

Although studies have clearly documented a negative relationship between family functioning and addiction, there should be more research to understand pathways through which addiction influences family functioning and vice versa. As described in this article, some of these pathways have been explored in prior research. However, there is need to expand the understanding of how specific risk factors within family relationships affect addictive behaviors and how addictive behaviors influence family dynamics. An example of this need is the lack of research on families in which multiple members exhibit addictions and how this dynamic affects family functioning and addictive behaviors over time.

Finally, additional research is needed to expand the understanding of the relationship between the family and addictions other than alcohol. The majority of family-based addiction research has examined the relationship between alcoholism and various forms of family functioning. Less research has examined the relationship between drug use disorders and family functioning, and there exists a relative paucity of research examining the association between family dynamics and other forms of addiction, including gambling or internet addiction. Therefore, research is needed to better understand how addictions other than alcohol are related to family functioning.

SEE ALSO

Drinking Patterns, Alcohol Consumption, and Aggressive Behavior, Alcohol's Effects on Sexual Arousal and Sexual Functioning, Alcohol and Sexual Violence

Glossary

Expressed emotion (EE) communication toward a family member that involves hostility, criticism, or being emotionally overinvolved.

Externalizing behavioral problems children's behaviors that are characterized by disruptive and often aggressive behaviors.

Internalizing behavioral problems children's behaviors that are most typified by dysregulation of emotion or mood.

Partner aggression Antagonistic verbal or physical behaviors toward a relationship partner.

Romantic dyad romantic relationship involving two partners.

Further Reading

Foran, H.M., O'Leary, K.D., 2008. Alcohol and intimate partner violence: a meta-analysis. Clinical Psychology Review 28, 1222–1234.

Harter, S.L., 2000. Psychosocial adjustment of adult children of alcoholics: a review of the recent empirical literature. Clinical Psychology Review 20, 311–337.

Homish, G.G., Leonard, K.E., Cornelius, J.R., 2008. Illicit drug use and marital satisfaction. Addictive Behaviors 33, 279–291.

Leadley, K., Clark, C.L., Caetano, R., 2000. Couples' drinking patterns, intimate partner violence, and alcohol-related partnership problems. Journal of Substance Abuse 11, 253–263.

Marmorstein, N.R., Iacona, W.G., McGue, M., 2009. Alcohol and illicit drug dependence among parents: associations with offspring externalizing disorders. Psychological Medicine 39, 149–155.

Marshal, M.P., 2003. For better or worse? The effects of alcohol use on marital functioning. Clinical Psychology Review 23, 959–997.

Mattson, R.E., O'Farrell, T.J., Monson, C.M., Panuzio, J., Taft, C.T., 2010. Female perpetrated dyadic aggression predicts relapse in a treatment sample of men with substance use disorders. Journal of Family Violence 25, 33–42.

Moore, T.M., Stuart, G.L., Meehan, J.C., Rhagitan, D.L., Hellmuth, J.C., Keen, S.M., 2008. Drug abuse and aggression between intimate partners: a meta-analysis. Clinical Psychology Review 28, 247–274.

O'Farrell, T.J., Birchler, G.R., 1987. Marital relationships of alcoholic, conflicted, and nonconflicted couples. Journal of Marital and Family Therapy 13, 259–274.

O'Farrell, T.J., Hooley, J., Fals-Stewart, W., Cutter, H.S.G., 1998. Expressed emotion and relapse of alcoholic patients. Journal of Consulting and Clinical Psychology 66, 744–752.

Rotunda, R.J., West, L., O'Farrell, T.J., 2004. Enabling behavior in a clinical sample of alcohol-dependent clients and their partners. Journal of Substance Abuse Treatment 26, 269–276.

Whisman, M.A., 2007. Marital distress and DSM-IV psychiatric disorders in a population-based national survey. Journal of Abnormal Psychology 116, 638–643.

Relevant Websites

http://www.al-anon.alateen.org/
http://www.alcoholfreechildren.org/
http://pubs.niaaa.nih.gov/publications/social/Module10JFamilies/Module10J.html
http://www.nida.nih.gov/parent-teacher.html

CHAPTER

32

The Intergenerational Transference of Addiction

Justine Campbell, Tian P.S. Oei
University of Queensland, St. Lucia, QLD, Australia

OUTLINE

Economic, mortality, injury, and menta l health statistics demonstrate that addiction is costly from an individual and societal level. Therefore, understanding the mechanisms by which such behaviors are transferred is an important component in creating treatment and prevention strategies for addictive behavior. It is clear from addictive behavior literature and research that the intergenerational transference of addiction has been well established. Estimated rates of alcoholism, smoking, illicit drug use, and gambling reveal that male and female children are more likely to have a father and/or mother who also demonstrate such behaviors compared to offspring in the general population. These results suggest that parents play a major role in the acquisition of addictive behavior in their children. Although evidence shows quite convincingly that the symptoms of addictive behaviors aggregate within families, it remains unclear that by what mechanisms these symptoms are transferred. More specifically, it is unknown whether the transmission of substance misuse or maladaptive behaviors is a direct causal relationship between parental and offspring behaviors, or due to other mediating variables such as genetic, environmental, or cognitive factors.

A number of reviews and studies have focused on the intergenerational "transfer of risk," whereby specific parental behaviors are associated with an increase in the possibility that similar or related problems will occur in the next generation, and have aimed to determine the processes by which this transfer occurs. Various theories have been proposed for the intergenerational transference of addiction; however, the existing dominant explanation for this phenomenon is genetic theory, which proposes that problematic behaviors are transferred from parent to child through biological processes. Nonetheless, this theory has resulted in a number of

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00032-2

unanswered questions suggesting that other influences may also play a part in the transference of the behavior. For example, environmental theorists suggest that influences such as family dysfunction and social class explain a large proportion of the risk. Yet, as with genetic theories, the extent that these factors influence the transmission of addictive behavior within families remains debatable, leaving a large proportion of variance in interfamilial transference unaccounted for.

More recently, cognitive theories of transference have become more prominent and argue that information regarding addiction and its subsequent effects are acquired during childhood, and that parents have a major impact on these acquisition processes and outcomes. In particular, cognitive theories propose that the observation of parental habits contributes to the child's beliefs and expectations of the subsequent effects, which in turn reinforces their future substance use or behavior. It is suggested that the acquired information is retained in the child's long-term memory and then triggered once their own associated behavior begins.

The familial resemblance of addiction has been demonstrated in innumerable family, twin, and adoption studies and has revealed interfamilial similarities for behaviors such as alcohol abuse/dependence; cigarette smoking; illicit substance use; and gambling between first-degree relatives. As such, this chapter will provide a brief overview of existing research for the intergenerational transference of addiction, and delineate the proposed mechanisms for this transmission within high-risk families. Finally, the chapter will conclude by discussing recent findings for the interfamilial transfer of cognitions as a mechanism for this phenomenon.

FAMILIAL SIMILARITIES FOR ADDICTIVE BEHAVIOR

Family Studies

The familial aggregation of addictive behavior has been demonstrated through numerous family, twin, and adoption studies and results from this research has shown that a family history of problematic behaviors is one of the best explanatory predictors of the initiation and maintenance of later problems in both clinical and community samples. For example, a family history of alcohol abuse/dependence has been shown to contribute almost half of the variance in their relative's problematic alcohol symptoms, which is further increased with the familial concentration of alcohol problems. That is, the more relatives with a history of alcohol problems, the greater the risk of their relatives displaying problematic drinking behaviors. In a study of over 8000 first-degree relatives of alcoholic probands, alcohol dependence was found to be two to three times greater than that reported by controls, with overall lifetime risk rates for alcohol dependence of 28.8% for relatives of probands, and 14.4% for controls. These children are also more likely to report coping, enhancement, and conformity among their motivations to drink, and are also likelier to drink alone, drink to induce intoxication, and drink due to the pleasurable taste of alcohol.

Tobacco smoking research has also revealed numerous familial similarities in behavior showing that offspring of persistent smokers have twice the smoking rates as offspring of nonsmokers, and that parental smoking predicts the initiation and quantity of the same behavior in their children. Offspring of parents who commenced their smoking behavior at an early age and are current persistent smokers have been shown to have the highest rate of smoking behavior (47%) compared to offspring of later onset/relapsing parents (30–34%), and abstaining parents (14%). These rates are consistent for parent and offspring cannabis use with first-degree relatives of cannabis abusing probands being almost six times more likely to abuse cannabis than relatives of control individuals. A moderate risk for the transference of gambling behavior also exists with a family history of gambling increasing the risk of offspring gambling behavior twofold (22–29%) compared to those whose parents did not report excessive gambling (9–15%). However, this risk is increased for children who have more than one family member with a gambling problem, and are over four times likelier to report having a gambling problem themselves, particularly if the family members are first-degree relatives.

This provides a strong evidence for the transference of addictive behaviors from parents to offspring and support for the theory that underlying mechanisms function within families to expedite this transference. However, it has been argued that the use of "at-risk" family studies fails to recognize the influence of genetic and environmental factors in the etiology of relative's addictive behavior compared with twin or adoptee studies where more observable genetic markers are available.

Twin and Adoption Studies

Twin studies, like family studies, have also shown that a family history of addiction is a consistent risk factor in developing alcohol dependence. These studies compare similarities between monozygotic (MZ) and dizygotic (DZ) twins of alcoholic parents raised together, and allow for investigations of genetic and environmental influences between individuals with 100% shared genes and those who share only 50% of

their genes. High heritability estimates of addictive behavior have been demonstrated in twin and adoption studies of adolescent tobacco use revealing estimates between 36 and 60%. However, varied results have been revealed for gambling addiction with heritability contributing between 35 and 54% of pathological gambling behavior, yet in contrast, meta-analyses of family and twin studies have only revealed a weak but significant heritability estimate of 16%. Various studies of alcohol abuse/dependence have demonstrated the familial nature of such disorders with MZ–DZ concordance ratios of approximately 2:1. Twin studies of male and female MZ and DZ twin pairs have consistently shown heritability scores for the likelihood of alcohol dependence symptoms ranging from 47 to 64%, with pair correlations for alcohol abuse or dependence (DSM-IV criteria) of 55 and 31% for MZ and DZ twins, respectively.

Gene-environment studies on adopted away children reveal an increased risk for alcohol and substance addiction in adoptees from alcoholic or substance using biological backgrounds (compared to control adoptees), whereas such abuse by the adoptive parents is not associated with greater risk of the same in the adopted child, suggesting a minor influence of environment. However, to further disentangle these results, the inclusion of stepfamilies with a nonalcoholic biological parent, and either an alcoholic or a nonalcoholic stepparent enables further scrutiny of specific hypotheses concerning genetic and environmental influences. Specifically, the inclusion of one nonalcoholic biological rearing parent with one nonbiological (either alcoholic or nonalcoholic) parent can in turn exclude the genetic transmission of alcohol problems in these offspring. As such, having an alcoholic biological parent significantly increases the risk of offspring developing alcohol abuse and dependence symptoms, with associations almost three fold greater for biological than adoptive and stepfamilies. However, results for nonbiological families are less straightforward such that having an alcoholic adoptive mother or an alcoholic stepfather predicts offspring alcohol abuse, therefore suggesting specific environmental transmissions. These heightened risks may therefore be attributed to other psychopathology in the biological parent, which may have contributed to the adopting out of the child, the termination of prior marriages, and/or some form of abuse or neglect by a stepparent.

The findings from twin and adoption studies indicate the importance of a genetic factor in the intergenerational transference of addictive behavior, yet like family studies, these imply that other mediating processes are functioning to initiate the transmission of such problems. Research assessing parent and child gender differences seem to support this suggestion, and imply that other mechanisms are at work rather than just direct genetic similarities or simple observation and imitation of parental role models.

Parent and Child Gender Differences

The familial relationships between parent and child addictive behavior appear to be moderated by parental and offspring gender; however, results appear to be varied when describing the relationship between same-sex and cross-sex parent/child dyads. Specifically, paternal addictions have been associated with greater levels of the same behavior in their sons, while maternal addictive behaviors predicted daughters increased behaviors suggesting a modeling or imitation effect of the same-sex parent. Yet equally, particular adverse behaviors in either parent has been shown to increase the risk of the same in their children; as does cross-sex effects where the parents' behavior has a greater impact on the opposite sex child's behavior; or more frequently, one parent has a stronger influence on both children irrespective of gender. Nonetheless, this implies that whereas addictive behavior symptoms aggregate within families, other processes are operating to mediate this intergenerational transference, and that the direct effects of parental gender and addictive behavior on offspring behavior are not so straightforward.

It has been argued that females are more sensitive to disruptions in their home environment, which may leave them vulnerable to addictive behavior, whereas males are more prone to genetic influences. For example, additive genetic factors have been found to account for approximately 60% of the variance in alcohol use in 17–18-years-old twin males, but only 10% in female twins; however, shared environmental factors accounted for 68% in female alcohol use, but only 23% in male alcohol use. It is therefore important to consider the role of certain social constructs transferred from the same-sex parent when discussing same- and cross-gender relationships in parent to child addiction transference, and how cognitions related to these constructs may be transferred and become significant antecedents in the initiation of problem behavior. Specifically, formation of gender identity and association with either masculine or feminine traits may play a significant part in the development of explicit gender-related styles of addictive behavior. For example, identification with masculine traits or exaggerations of conventional female stereotypes may contribute to the risk of addiction. As such, the motivations to use substances or gamble may differ between sexes, but underlie the transference of cognitions associated with typical gender roles from parents to children. This issue will be explored more thoroughly in a later section.

It is therefore implied that the pressure to conform to traditional gender roles may leave females more vulnerable to maladaptive outcomes due to family dysfunction, whereby their propensity to display interpersonal connectedness and concern for the welfare of others make them more emotionally affected by family discord.

Summary and Critique

Family, twin, and adoption studies have provided strong evidence that the symptoms of addiction aggregates in families, and the intergenerational transference of these problems can be attributed, at least in part, to a genetic mechanism. Many of these analyses have reported heritability statistics of 40 to 60%; with familial concentration of addiction contributing to elevated levels of symptom transference. However, various problems arise insofar that "at-risk" family studies often confound biology and environment such that most families share these elements. The development of twin studies reduced this limitation, arguing that greater similarities would occur between identical twins that share 100% of their genetic material, compared to fraternal twins, who only share 50%. The heritability rates for these studies provide further support for a genetic influence in the transference of addictive behavior, with a twofold risk of problems in identical twins, compared to nonidentical twins. Similarly, adoption and stepfamily studies have revealed that addictive behaviors in the biological parent, but not the nonbiological parents, are frequently associated with a greater risk of similar problems in the adopted or stepchild. Furthermore, the examination of the moderating roles of parent and child gender has indicated that the genetic load for addiction is equivalent in both genders; however, it would appear that conformity to social roles contributes to particular styles of addictive behaviors, particularly in females.

Nonetheless, twin and adoption studies are not without their limitations (although the evidence is mixed), particularly in regards to assortative mating and equal environment assumptions for twin pairs, and pre/post natal influences or selective placement for adoptees. In a meta-analysis of 50 family, twin, and adoption studies of substance use, heritability estimates much lower than those rates normally cited in the literature were revealed. In fact, after restricting those studies that contribute highly to the genetic hypothesis, the heritability scores for the interfamilial transference did not exceed 26%. These findings indicate that up to 70% of the variance in addictive transference is attributable to other factors. Numerous studies have explored various biological, environmental, and cognitive factors to explain the processes by which the transference of alcohol problems is attributable.

MECHANISMS IN THE INTERGENERATIONAL TRANSFERENCE OF ALCOHOL PROBLEMS

Human Genome Studies and Endophenotypes

Family studies have indicated that the human genome influences the risk of developing certain addictions, which has led to a plethora of research examining the effect of specific chromosomal regions and genes to further explain the mechanism for the intergenerational transference of such disorders. These include both linkage and association studies. Linkage studies involve families with multiple affected individuals, and identify variations within segments of DNA and chromosomal regions that are common among family members. The linkage concept suggests that genes located in close proximity to each other are likelier to be inherited together from one parent, than two distal genes. In contrast, association studies are not restricted to family samples and can use unrelated controls to assess the relationship between specific genes and a particular outcome across families. Mixed linkage and association findings of alcohol dependence, substance use, tobacco smoking, and gambling have implicated, among others, the μ-opioid receptor gene; the serotonin transporter gene; the dopamine receptor gene; and GABA receptor genes. However, much of this research has yielded inconsistent results across groups of varying sample size, ethnicity, and clinical diagnoses. Furthermore, given the exhaustive list of proposed genetic markers, restricting the genetic transmission of addictive behavior within families remains indistinct. Nonetheless, despite some ambiguity in the results, the findings from human genome studies support the suggestion that genes operate as a mechanism for the interfamilial transference of addiction, such that similarities in genetic makeup between parents and offspring contribute to similarities in addictive behavior.

Similarly, it is argued that an endophenotype for a disorder should be heritable if there is a direct relationship between it and other susceptibility genes for the disorder. One genetically influenced phenotype that has been associated with increased risk for substance use addiction is the reduction of P300 (P3) amplitude during event-related potential recordings. The P3 is a brain potential which indicates the amount of attentional resources required for encoding new information in working memory. The reduction in this P3 wavelength has shown to be more similar in MZ than DZ alcoholic; however, heritability estimates appear to vary across gender, with percentages around 65% for boys, and 35% for girls, who were also substantially affected by shared environment. A parental transfer risk has been found in children of alcoholic parents

exhibiting reduced P3 amplitude; however, this parental risk effect appears greater in high-risk families, compared to low-risk families who were less likely to be exposed to the negative effects of alcohol or other substance abuse.

Summary and Critique

A plethora of research has provided mixed evidence for a specific gene, or genes, to explain the mechanism of the intergenerational transference of addiction. Similarly, whereas the identification of observable endophenotypes has provided similar heritably estimates as family studies, it is argued that given the polygenic nature of most genetic influences, it is not possible to implicate a specific gene, or gene combination, that contributes to the vulnerability toward specific addictions. Nevertheless, these studies provide some evidence that the role of genetics is a likely contributor to the transference of problematic behavior within families. It is argued however that the impact of biological influences can only be understood when evaluated in the context of environmental contributors, given that these factors have also shown to contribute to a family history of addictive behavior. In other words, the interaction of genetic and environmental influences may increase the risk of children of parents with specific addictions developing similar disorders under certain family environmental circumstances.

GENOTYPE-ENVIRONMENTAL THEORIES

Genotype × Environment (G×E) research that has assessed lifetime prevalence of problematic addictive behaviors has noted that individuals at high genetic risk for substance use, alcoholism, and gambling are usually also exposed to high-risk environments. As such, the intergenerational mechanisms that occur remain ambiguous such that the increased rates of problematic behaviors observed in children of problem drinkers, smokers, substance users, or gamblers do not distinguish whether it is the genetic transference from parents to offspring, poor family functioning due to problematic parental behavior, or a combination of genetic and environmental influences which contribute to the increase in risk in offspring addiction. To explore this, studies assessing the intergenerational effects of addictive behavior often discriminate between those environmental factors that are shared by siblings (SE or *c*), which include parental and family influences, and nonshared environmental factors (NSE or *e*) which are unique to siblings (e.g. peer groups), and the genetic component shared by parents, offspring, and siblings (G or *a*). The majority of research into the transference of addiction within families normally focuses on the first and last of these factors in determining the contribution of genetic and environmental influences. Estimates of the percentage of variance in addictions due to shared environment vary excessively among the literature. For example, several studies assessing shared and unique environmental factors on cannabis use found evidence for shared influences with estimates ranging from 26 to 85%, whereas estimates for cannabis abuse/dependence ranged between 3 and 29%.

Many G×E studies have found however that common environment interacts with heritability estimates according to age, zygosity, and gender. Assorted studies of alcohol transference have reported shared environment estimates between 37 and 71% among adolescent and young adults. Alternatively however, in a large twin study of MZ and DZ male twin pairs, common or shared environment only accounted for 3–11% of the variation in the development of alcohol abuse and dependence, with no major G×E effects.

However, in stark contrast before and after controlling for the effects of sociodemographic and psychiatric predictors on the risk of alcohol dependence, a more recent study found that shared environment did not provide any variance to the disorder. It was conjectured, though this finding did not exclude possible G×E interactions given that the interaction between genetic and shared environmental effects contribute to similarities of biological siblings reared together, but not to unrelated siblings reared together, or to biological siblings reared apart.

It would appear by these results that, as a mechanism, shared environmental factors add little to the explanation of the intergenerational transference of addictive behaviors, with some studies leaving over 90% of the variance in addiction transference in families unaccounted for. Additionally, it would appear that G×E interactions appear greater for unshared environmental factors, suggesting that indirect effects of parental addiction act as a mechanism for the interfamilial transmission of addictive behaviors.

Family Dysfunction

Problematic addictions adversely affect both the global family environment and the psychological well-being of the addicted parent's offspring. These difficulties with both family and personal functioning have resulted in children of addicted parents being recognized as an at-risk population for similar addictions, either directly as a means of coping, or indirectly through the association with deviant peers due to low parental monitoring. However, the process of transference of addictive behaviors from parent to child due to

family dysfunction remains somewhat unclear given that negative family processes may either be the cause or the result of problematic parental behaviors. For instance, in a German alcohol study it was found that a family history of alcoholism was unrelated to adolescent alcohol problems. However, the parental co-occurring psychiatric diagnoses, and the offspring's perceived parental rejection was related to the child's aggression/delinquency, which in turn led to association with substance using peers and thus, to alcohol problems. When those children diagnosed with conduct disorder and/or antisocial personality disorder were removed from the analyses, however, aggression/delinquency was no longer associated with peer group substance use, and therefore no longer acted as a predictor for alcohol use problems. Family psychopathology however remained a significant predictor of alcohol problems suggesting that this factor increases the risk for a range of behavioral and emotional problems in children of alcoholics. It should be noted however that some substance use studies have found no difference in the specificity of problematic family environments between children of substance users and other children exposed to other family stressors such as parental divorce, death, or major illness. This argues against the notion that parental addictive behavior is a sufficient cause for family dysfunction, and that unstable family environment appears to be equally prevalent among families exposed to other significant stressors. Irrespectively, it appears that a dysfunctional family environment operates as a partial mechanism for the intergenerational transference of addiction; however, what remains unclear is whether this process operates directly or indirectly.

Summary and Critique

As shown, a large amount of literature has addressed the relationship between genetic and environmental factors in describing the mechanism for the intergenerational transference of addiction, yet findings remain inconclusive. Specifically, genetic and environment studies have provided estimates from 0 to 71% for shared environment, and inconsistencies remain as to the G×E interaction. Furthermore, the discrepancies in the findings assessing the relationship between varied elements of family dysfunction and the interfamilial transference of addiction seem to suggest that addictive behavior in offspring may operate through other influences resulting from problematic parental addiction.

Overall, the range of biosocial theories have provided some evidence (albeit mixed) for heritable and milieu related influences in the intergenerational transference of addiction. However, this literature leaves a large proportion of the variance in this process unaccounted for. Further, studies show that certain children do not develop addiction disorders despite having a parent with the same disorder, whereas others acquire these regardless of nonaddicted parents. It is therefore suggested that other mechanisms are operating which account for a part of the variance in interfamilial addictive behavior, which may serve as either protective or risk factors. Particularly, it is proposed that specific social learning cognitions regarding the anticipation, expectancy, memory, and modeling of certain behaviors are fundamental in determining addiction in children and young adults.

COGNITIVE PERSPECTIVES

The role of social factors cannot be overlooked as significant antecedents in the initiation of addictive behaviors. However, certain social constructs (e.g. social norms and gender identification) need to be considered simultaneously with the cognitive aspects associated with them and their interactive effects on the intergenerational transference of these problematic behaviors. Specifically, children develop early expectations about gender through observation of parents and then form schemas that influence how they perceive the behaviors of men and women, and in turn, how these behaviors adhere to social norms, most particularly with alcohol consumption, cigarette smoking, and substance use.

Given that gender plays a fundamental role in the formation of identity (and social norms), association with either masculine or feminine traits may be a significant means through which distinctive gender-related styles of substance use or gambling behavior develop. For instance, males typically display a considerably higher prevalence of gambling, drug and alcohol related problems than women, and identification with the masculine role such as dominance, assertiveness, and independence may contribute to this. Alternatively, exaggerations of conventional female stereotypes may encourage females to be submissive, helpless and dependant thus increasing the risk of their addictive behaviors. It should be noted however that convergence in gender behaviors and changes in societal norms related to gender identity has led to similarities in substance use quantities or gambling frequency in both males and females. This would indicate that the changes in perceptions of gender role differentiation may also be mirrored in changes of addictive behavior. As such, the motivations to participate in substance use or gambling (e.g. coping with negative emotions, stress control, social facilitation, adherence to perceived societal norms) may be different in each sex but underlie the transference of cognitions associated with typical gender roles from parents to children.

The increase in the number of young adolescents and adults participation in gambling or risky substance use suggests that antecedent characteristics may exist prior to the individual's first experience with substances or gambling. For instance, when asked to select a small number of lottery tickets from a larger selection, primary school aged children were more likely to utilize random selection and rationalization for selection using reference to favorite, lucky, or important numbers suggesting the use of strategic methods when gambling. Similarly, children as young as 6 years old have shown to possess some understanding of the contextual, motivational, and normative aspects of alcohol consumption and behavior. One study has revealed that by the age of three most children could identify alcohol type by photographs, and that these children held two common alcohol schemas also held by the greater culture: (1) that alcohol consumption is done more by adults than children and (2) that alcohol consumption is done more by males than females.

Based on these findings, it has been argued that a large proportion of intra-family addictive behaviors may occur through the transference of associated cognitions. Previous alcohol and tobacco smoking research seems to support this possibility and indicates that adolescent's perception of their parents' approval or disapproval of their substance use was related to the child's drinking or smoking habits, such that perceived parental disapproval was associated with increased self-efficacy for refusing alcohol and/or cigarettes, and lower consumption.

EXPECTANCIES AND REFUSAL SELF-EFFICACY

Oei and Baldwin have proposed a cognitive theory to explain the development and maintenance of problem drinking within an individual. It is argued that this model can be extended to explain the relationship between parental addiction and related offspring behavior. For example, this model examines the relationship between Alcohol Expectancies (AE) and Drinking Refusal Self-Efficacy, which are beliefs about specific outcomes and self-control behavior associated with alcohol consumption. The theory proposes a two-process model of alcohol use and abuse: (1) an acquisition phase, based on instrumental learning or modeling processes in which AE are formed, and (2) a maintenance phase based on classical conditioning, in which unconscious conditioned processes automatically initiates a drinking response. In terms of an interfamilial transference of addiction, expectancies as to the effect of a particular addictive behavior originate as a result of parental modeling, and once these cognitions become established they guide the child's behavior when exposed to the factor (i.e. alcohol, substances, gambling). Once initiated these expectancies are reinforced, thus leading to the maintenance of addictive behavior.

Expectancies are generally expressed in the form of contingencies, or "if…then" statements (e.g. If I drink alcohol/use substances/gamble etc… then I will be happy/sad/rich/poor). These expectancies predetermine the individual's choice to undertake (or not undertake) specific behaviors and also their subsequent actions, which is driven by the anticipation of the ensuing effects such as increased sociability or tension reduction. Successive confirmation of these expectancies can reinforce the behavior and, in the case of optimistic outcomes, place the individual at risk of persistent problem behavior. In contrast, self-efficacy has been defined as the perceived ability to refuse a substance or participate in a behavior in a specific situation (e.g. with friends; feeling sad), rather than whether or not one chooses to commit this behavior. Findings from previous research would suggest that the constructs of expectancies and self-efficacy form the foundations of the transfer of addictive behavior between generations, and have repeatedly been demonstrated in alcohol, smoking, and gambling research.

THE INTERGENERATIONAL TRANSFERENCE OF ADDICTION RELATED COGNITIONS

The formation of cognitions prior to actual experience indicates that the knowledge forming these beliefs is derived from sources other than actual addictive behavior. For children, arguably the most accessible and significant models displaying such behaviors are their parents. It has been suggested in previous literature that exposure to parental behavior influences the child's behavior directly through imitation and modeling; however, adoption studies on substance use transference suggest that risk for intergenerational transmission is elevated even in the absence of contact with the substance using parent, making the argument for observation as a requirement for transmission debatable.

It has been suggested that children with a family history of addictive behaviors differ from those without a family history in terms of their cognitive expectations regarding the particular substance or behavior's effects. Research assessing the intergenerational transference of gambling related cognitions has revealed that parental gambling beliefs were directly related to the same in their children, which subsequently predicted their offspring's gambling behavior. Additionally,

expectancies have been shown to differ between offspring of alcoholic parents and offspring from nonalcoholic families, such that adolescents with a family history of alcohol misuse expected more enhanced cognitive and motor abilities. This would imply that children adopt their parents' expectancies regarding the perceived advantageous effects of certain behavior, without the influence of their own similar experiences. For example, it has been proposed that children may acquire the expectation of enhanced functioning via their alcohol abusing parents' self-report of improved performance while drinking, or via the child's observation of the parents' reduction in withdrawal symptoms, once alcohol consumption is resumed. This argument has been supported by the finding that offspring at high risk of developing alcoholism (i.e. presence of alcoholic father and uncle) held beliefs about the effects of alcohol that were similar to their parents, compared to low-risk controls.

Less is known however about the transference of self-efficacy beliefs between parents and children with very sparse research being conducted to investigate this occurrence. Some research that may provide tentative evidence for the transference of addiction related self-efficacy beliefs are studies exploring the effect of self-fulfilling prophecies on children's alcohol use behavior. These studies revealed that parents' beliefs about their offspring's drinking self-efficacy were predictive of their children's alcohol use behavior such that unfavorable beliefs significantly predicted their children's increased alcohol consumption. Similarly, other substance use research has revealed that parental smoking self-efficacy and beliefs about youth smoking were related to both theirs and their children's cigarette smoking behavior. These studies found that parents reporting lower smoking self-efficacy and higher pro-social youth smoking beliefs tended to have children who were regular smokers. Whereas these studies provide some speculative indication that parental self-efficacy beliefs may transfer within families, they do indicate that parents' beliefs about substance use may at least indirectly contribute to their children's behavior.

What appears to be more likely however is that whereas observation of parents' addictive behavior appears to initiate the development of expectancies in children prior to their first experience with a particular substance or behavior, it is the understanding and adoption of similar related beliefs and self-efficacy as their parents that maintains children's actions as adolescents and young adults. Specifically, whereas parents' expectancies do indeed contribute to children's cognitions, they play a greater role in impacting their offspring's personal self-related beliefs regarding their substance use or behavior, as opposed to their beliefs about the effects or outcome of the behavior itself. This suggests an interaction between children's awareness of their parents' expectancies, and their parent's behavior associated with these expectancies (see Fig. 32.1).

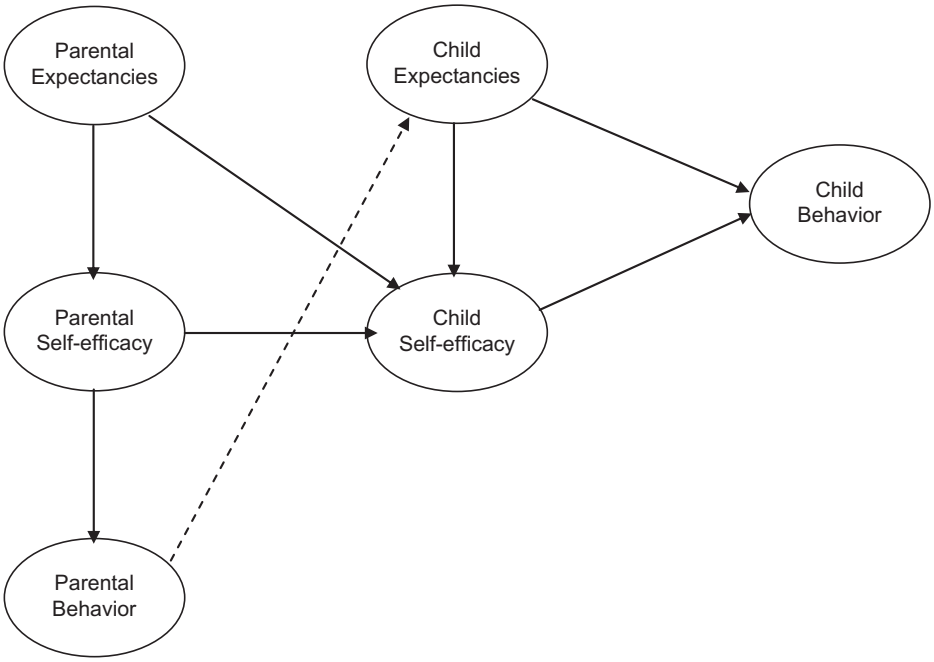


FIGURE 32.1 The intergenerational transference of addiction cognitions indicating the early development of expectancies via the observation of parental behavior (dotted line) and the maintenance of the behavior via reinforcement of expectancies and the understanding and adoption of similar parental beliefs and expectancies (full lines).

This emphasizes the importance of appropriate parental communication regarding the use of substances or gambling on their children's ongoing behavior. This is evidenced by the fact that parental communications regarding appropriate lawful substance use or gambling behavior are more likely to be internalized by their offspring within amiable and supportive parent–child relationships, whereas familial problematic behaviors have been shown to be related to poor family communication. As such, the independent roles of expectancies and refusal self-efficacy have great implications for the prevention and treatment of problematic and addictive behaviors rather than one overarching remedy. Maladaptive cognitions regarding positive outcomes of excessive substance use or gambling could be challenged and altered prior to the onset of the child's first experience with the behavior, whereas recognition of lower self-efficacy beliefs could be managed through familiarization and understanding of cue states, and the development of strategies for avoiding specific situations in the future. Whereas this suggestion may not be revolutionary, the integration of addiction prevention and treatment into a multi-systemic, family-oriented treatment and prevention therapy has the capacity to deal with tertiary problems in parents while simultaneously concentrating on primary or secondary prevention with their offspring. Strategies such as educating parents as to their role in the transference of addiction related messages to their offspring may reduce children's expectancies and self-efficacy beliefs prior to the onset of their first substance use or gambling experience.

CONCLUDING COMMENTS

The contributions of research to the "transfer of risk" literature regarding addictive behaviors within families has provided further understanding of the extent in which our cognitions impact upon others beliefs and behavior, specifically when the individual is vulnerable to suggestion and likely to adopt similar beliefs and behavior merely through observation. The findings also suggest that the development and maintenance of a specific behavior does not unconsciously occur as a result of direct mimicry or imitation, nor does this behavior remain uninfluenced by these same cognitive processes throughout our lifetime.

What is more likely, and demonstrated in the model, is that the development and transfer of addiction requires some aspect of an observable behavior. That is, although such communication expresses parent's beliefs related to the addiction, it is argued that it is the observation of the associated behavior that establishes these cognitions in their offspring. This indicates that parental behavior is important in the initiation of their younger offspring's related expectancies; however, once these cognitions are established, it is parental cognitions that become more prevalent in the maintenance of their child's cognitions and resultant addictive behavior throughout adolescence and young adulthood.

Nonetheless, it is proposed that the current argument has contributed to current literature by extending and strengthening an already well-developed theoretical framework to examine the complex relationships between parent and child addictive behaviors, and their role in predicting future problematic behavior. The challenge for future research is to determine the applicability of this model to multiple generations and to what extent this impact has on more problematic addictions, and even more importantly, the contributions that this theory has on prevention and treatment outcomes.

SEE ALSO

An Evolutionary Perspective on Addiction, Cognitive Factors in Addictive Processes, Families and Addiction

List of Abbreviations

AE	alcohol expectancies
DZ	dizygotic
G×E	Genotype × Environment
MZ	monozygotic

Further Reading

Baldwin, A.R., Oei, T.P., Young, R., 1993. To drink or not to drink: the differential role of alcohol expectancies and drinking refusal self-efficacy in quantity and frequency of alcohol consumption. Cognitive Therapy and Research 17, 511–530.

Bandura, A., 1977. Self-efficacy: toward a unifying theory of behavioral change. Psychological Review 84, 191–215.

Bandura, A., 1982. Self-efficacy mechanism in human agency. American Psychologist 37, 122–147.

Campbell, J.M., Oei, T.P.S., 2010. A cognitive model for the intergenerational transference of alcohol use behavior. Addictive Behaviors 35, 73–83.

Dowling, N.A., Jackson, A.C., Thomas, S.A., Frydenberg, E., 2010. Children at Risk of Developing Problem Gambling. The Problem Gambling Research and Treatment Centre, University of Melbourne and Monash University, Melbourne.

Hasking, P.A., Oei, T.P.S., 2002. The differential role of alcohol expectancies, drinking refusal self-efficacy and coping resources in predicting alcohol consumption in community and clinical samples. Addiction Research and Theory 10, 465–494.

Oei, T.P.S., Baldwin, A.R., 1994. Expectancy theory: a two-process model of alcohol use and abuse. Journal of Studies on Alcohol 55, 525–534.

Oei, T.P.S., Burrow, T., 2000. Alcohol expectancy and drinking refusal self-efficacy: a test of specificity theory. Addictive Behavior 25, 499–507.

Relevant Websites

http://australia.gov.au/topics/health-and-safety – Australian Government Strategies.
www.abc.net.au/news – Australian Broadcasting Corporation.
www.abc.net.au/health – Australian Broadcasting Corporation – Health and Wellbeing.
www.drinkwise.org.au – DrinkWise Australia.
http://www.alcohol.gov.au – National Alcohol Strategy.

CHAPTER

33

Peer Influences on Addiction

Clayton Neighbors [*], *Dawn W. Foster* [*], *Nicole Fossos* [$]

[*] University of Houston, Houston, TX, USA, [$] University of Washington, Seattle, WA, USA

OUTLINE

OVERVIEW

Peer influences, and social influences more generally, are among the most powerful determinants of behavior. With respect to addiction, peer influences seem to be most salient during the period from adolescence to young adulthood and appear to diminish as individuals become more oriented toward work and family. This does not mean that peer influences are no longer important, but that they tend to be less salient and less likely to be focused around addictive behaviors. There are certainly exceptions to the typical gradually decreasing influence of peers on addiction, but for the purposes of this article we refer to processes that have been most often studied among adolescents and young adults. The article begins with a brief review of several theories related to peer influence. The review of theories is intended to provide context and ways to synthesize thinking about peer influence. Next, we discuss direct and indirect influences on addictive behaviors. Direct peer influences range from polite requests to overt commands. Indirect peer influences are more subtle, including modeling (this involves observation and imitation) and perceived norms (subjective views on what is normal). Finally, we discuss resistance to peer influence and conclude with a brief summary.

THEORIES OF PEER INFLUENCE

In this section, we briefly summarize the major theories of peer influence and consider their relevance to alcohol and drug use. Several social influence theories have not been extensively examined in the context of peer influences on addiction but are clearly relevant. Theories of social influence have considerable overlap with each other and there are relatively few contradictions among theories in this domain. Rather, considering the different theories of social influence seems

Principles of Addiction, Volume 1
DOI: http://dx.doi.org/10.1016/B978-0-12-398336-7.00033-4

analogous to viewing an object from different angles and vantages. Certain attributes are more prominent from some perspectives versus others, whereas some attributes are observable only from a particular vantage. We believe that familiarity with multiple theories of social influence is likely to provide the broadest and most comprehensive theoretical understanding of peer influences on addiction. Thus, we provide brief overviews, with more or less detail of multiple theories, which vary in the extent to which they have been empirically applied to addiction.

Social Comparison Theory

Social Comparison Theory describes the causes and consequences of comparing oneself to others. The theory was proposed by Leon Festinger in 1954 as a formal theory with nine hypotheses and eight corollaries. It has remained a dominant theory of social influence and the central theories related to social influence. An initial presumption of this theory was that people compare themselves with others because they want to gauge themselves accurately, especially when objective criteria are not available. Thus, we look to others to help us determine how we should behave. Subsequent research has revealed multiple motivations and functions of social comparisons. Social comparisons can be used to bolster self-esteem or justify behavior (e.g. "I may drink and smoke pot every now and then but at least I'm not like Craig, who can't go a whole day without getting high."). Social comparisons can also function as aspirations in comparisons with those whom we wish to be like (e.g. "I wish I could drink as much as Suzanne without getting totally obliterated," or, "I wish I could quit smoking like Marvin did."). Other motivations include wanting to form a common bond with others, altruism, and self-destruction. The impact of any given social comparison depends on a number of factors including the motivation for making the comparison, relationship to the comparison target, centrality of the comparison dimension to one's sense of self, and direction of comparison. Comparisons with others whom we perceive as being better than we are on a given comparison dimension tend to make us feel worse about ourselves unless we see the comparison target as a source of inspiration. A comparison suggesting that everyone else seems to have fun when they are drinking but I get out of hand and do things I feel embarrassed about later is likely to result in negative affect and self-admonishment. In contrast, an AA newcomer getting a 30 day sobriety chip may be inspired by a fellow member receiving a 6-month sobriety chip. Comparisons can also be made in the opposite direction with others who are worse than we are on some dimension. These types of comparisons can make us feel better about ourselves unless we empathize with the comparison target. A comparison such as "I may drink a lot, but at least I don't drink as much as Nate" may provide justification and affirmation of one's own drinking behavior. On the other hand, comparing ourselves to a friend, with whom we have regularly used substances, who just got busted, may not make us feel so good. Another dimension of social comparison that has been extensively examined is the choice of comparison targets. People may choose comparison targets based on their ability to inspire us or to make us feel better about ourselves. Consistent with social identity (see below), comparisons to others with whom we most strongly identify often have the greatest influences on our behavior. Social comparisons with friends or coworkers will likely have more influence on our behaviors than comparisons with strangers.

Theory of Planned Behavior/Reasoned Action

Theory of Planned Behavior/Reasoned Action was proposed by Ajzen and Fishbein and suggests that behavior is determined by intentions, attitudes (beliefs about a behavior), and subjective norms (beliefs about others' attitudes toward a behavior). The theory was later expanded to the Theory of Planned Behavior wherein perceived behavioral control (beliefs about one's ability to perform a behavior) and behavioral intentions predict behavior. More extensive consideration of these theories are provided elsewhere in this encyclopedia, but the construct of subjective norms is worth additional consideration here. The construct of subjective norms is analogous to the injunctive norm component of other social norms theories and is one operationalization of the more general construct of social norms. Subjective norms refer specifically to how others whom we care about would feel about us engaging in a particular behavior (e.g. "Most people that are important to me think that I should not smoke cigarettes."). According to the theories of reasoned action and planned behavior, if I believe that important people in my life would disapprove if I smoked cigarettes, I should be less likely to intend to smoke cigarettes and subsequently less likely to actually smoke cigarettes. Subjective norms are distinct from other operationalizations of social norms in two ways. They focus exclusively on important others as the reference group and the behavior of relevance is on the perceiver's behavior rather than the behavior in general. Thus, the question is not about the extent to which I think others about whom I care approve or disapprove of smoking but rather the extent to which I think others about whom I care approve or disapprove of MY smoking. This may seem like a small and subtle distinction, and in many cases it probably is. However, in other cases it

may be an important distinction. For example, parents with moderate or favorable views on legalization of marijuana may be less approving of marijuana use by their teenage daughter. While subjective norms and injunctive norms share overlapping features, injunctive norms are framed such that they are inherently consistent with Social Identity Theory and Social Impact Theory (see below) and do not include the stipulations regarding the reference group (important others) or the specification that the behavior in question be the person's own behavior.

Social Identity Theory

Social Identity Theory considers how group membership is incorporated into our self-concept and how this affects our views of other members and nonmembers of our groups as well as members of rival groups. The theory has been used extensively in considering in-group favoritism and out-group discrimination but has received only minimal attention with respect to substance use. Specifically, social identity has been found to moderate the influence of others on substance use. For example, perceptions of the prevalence and approval of drinking among other members of one's group have been found to be more strongly associated with one's own drinking when the individual identifies more closely with the group. Identification with others augments their influence on us. In considering social groups among high school students, an athlete should be more influenced by other athletes, whereas drama club students should be more influenced by other drama club students. Identification with other substance users is highly correlated with personal substance use behavior. This is likely due in part to selection (substance users seek affiliation with other substance users) and socialization (affiliation with other substance users influences substance use). In sum, the people with whom we most strongly identify have the largest influence on our behavior.

Social Impact Theory

Social Impact Theory proposes three variables which are presumed to predict the degree of social influence that a particular source may have. According to this theory, social influence is the product of the strength, immediacy, and number of persuasive agents. Here, strength refers to how much we care about the person or persons who may be attempting to influence us or with whom we are comparing our behavior. This strength dimension is similar to social identity and its inclusion as a predictor of social influence is consistent with most theories of social influence. Immediacy and number are relatively unique factors proposed by the Social Impact Theory. Immediacy refers to proximity of the source or sources of influence. Potential sources of influence are more powerful when they are more proximal than when they are more distal. For example, according to this theory, a teen is likely to have greater difficulty resisting peer influence to smoke a joint in person (e.g. friend pulls out a joint and says, "Look what I have. Do you want to smoke it with me?") than the same request by email or by phone text message. Number refers to the number of independent sources of social influence. More sources of social influence are more persuasive than fewer. Thus, according to the theory, all other things being equal, three friends encouraging a young adult to do tequila shots would be more persuasive than one friend. Basic research on Social Impact Theory suggests that the influence of number diminishes at about four or five independent sources. In sum, Social Impact Theory can be viewed as a theoretical extension of Social Comparison and Social Identity Theories with a unique focus on the proximity and number of sources of influence.

Prototype/Willingness Model

Prototype/Willingness Model is an extension of the Theory of Reasoned Action and posits two paths, a reasoned path and a social reaction path, to engaging in risky behaviors such as substance use. The reasoned path represents an intentional style of processing whereby actions are premeditated and are a function of behavioral intentions. In turn, intentions to engage in a behavior are influenced by one's attitudes and perceptions of other's attitudes toward the behavior (i.e. subjective norms). In contrast, the social reaction path represents a heuristic-based style of processing suggesting that there are times when behavior is unintended and occurs in situations that facilitate risky behaviors such as substance use. The theory suggests that in some risky situations, it is not reasoned decision making (behavioral intentions), but behavioral willingness that determines decisions to engage in substance use. For example, an adolescent may have no predetermined intentions to smoke marijuana but may be perfectly willing to, should an opportunity or situation arise. Similarly, an individual may have no intentions of using a hard drug (e.g. crack cocaine or methamphetamine) but be willing to try it if the occasion arises. Risk prototypes refer to our impressions of individuals who engage in particular behaviors and are presumed to predict behavioral willingness. For example, an individual with a favorable prototype of recreational cocaine users would presumably be more willing to try cocaine in a situation where it becomes available than someone with a less-favorable prototype.

One of the assumptions underlying this theory is that while a person chooses to engage in substance use, substance use often is not planned or even intended. A second assumption is that individuals have distinct social images/representations (prototypes) of the type of person their age who engages in substance use (typology of a person rather than description). These prototypes are associated with behavioral willingness such that the more favorable the image of a substance user, the more likely they are to engage in substance use and to accept the social consequences of being seen as a substance user. An additional assumption of the Prototype/Willingness Model is that the relationship between intentions and behavior becomes stronger than the relationship between willingness and behavior over time.

Problem Behavior Theory and Peer Cluster Theory

Problem Behavior Theory and Peer Cluster theories have both had significant impact on our understanding of adolescent substance use and have been influential in shaping more recent perspectives. Problem Behavior Theory was originally proposed by Richard Jessor as a way of predicting proneness to deviance (unconventionality) and suggests multiple variables that instigate or prevent problem behaviors, and the balance between these variables predict engagement in problem behavior. Variables which instigate deviance have been defined as risk factors, whereas those which prevent or reduce the likelihood of problematic behaviors have been defined as protective factors. Problem behaviors are defined by age and societal norms. For example, consuming two beers at the age of 10 is conceptualized as problem behavior and a social norms transgression; however, consuming two beers at the age of 26 is typically not considered problematic. Deviance proneness leads to a wide range of problem behaviors, and these behaviors positively relate to each other and negatively correlate with conventional or prosocial behaviors. In addition, deviant behaviors in youth are a predictor of behavior problems in adulthood.

Peer Cluster Theory proposes that engagement in behaviors such as substance use is supported or discouraged by interactions with peers. The peer cluster is distinct from the peer group and is defined as a small unit of peers that shape attitudes and influence behaviors to a greater extent than other social network factors. The peer cluster has a powerful effect on an individual and interventions targeting risky behaviors such as addiction to illicit drugs are effective when the individual is separated from the peer cluster or the intervention targets the peer cluster directly.

Deviance Regulation Theory

Deviance Regulation Theory is a recent addition to the literature related to peer influence and has been evaluated only in a handful of studies. The theory provides compelling arguments suggesting that individuals are influenced more by deviance from the norm, rather than by conformity to others. That is, behavior is most influenced by our desire to stand out (deviate) from others in positive ways and not to stand out in negative ways. This is a novel perspective in contrast to most theories of peer influence which implicitly assume that individuals desire to actively conform or to fit in with their peers. The perceived attitudes or subjective norms of the relevant reference group and one's own attitudes, in combination with the situational context, determine what counts as deviating in positive versus negative ways. For example, being the only nondrinker at a fraternity keg party may represent negative distinction, whereas consuming the largest quantity of cheap beer while doing a keg stand at the same party may constitute positive distinction. What is viewed as positive and negative is relative to the individual, reference group, and situation. The theory has novel implications regarding message framing for peer influence based prevention strategies.

Messages that target behaviors viewed as common and healthy should emphasize the consequences of not engaging in the healthy behavior. Messages that target behaviors viewed as common and unhealthy should emphasize the benefits of not engaging in the unhealthy behavior. Messages that target behaviors viewed as uncommon and healthy should emphasize the benefits of engaging in the healthy behavior. Finally, messages that target behaviors viewed as uncommon and unhealthy should emphasize the consequences of engaging in the unhealthy behavior. In each case, the theory proposes that the key is to focus on deviance, rather than conformity, with praise for positive distinction and reproach for negative distinction. The theory has much intuitive appeal and may help explain failures of some public campaigns targeting addictive behaviors.

Social Learning Theory and Social Cognitive Theory

Social Learning Theory applies to several human behavior theories in which the acquisition and maintenance of behaviors such as addictive behaviors depend on the connections between personal factors, environmental factors, and the behavior. Social Learning/Cognitive Theory, to which Albert Bandura greatly contributed, focuses on several key constructs including differential reinforcement, vicarious learning, cognitive processes, and reciprocal determinism.

Differential reinforcement takes place when a behavior results in positive or negative consequences received from the environment or the self. This helps explain why behaviors may change with the environment. Note that consequences to behavior are often social consequences. For example, a teenager using cocaine with peers at a party may receive social approval; however, the same behavior, if observed or discovered by parents or other authorities would likely result in strong disapproval and additional unwanted consequences for the teenager. The likelihood that this teenager will engage in cocaine use is greater if he or she has a positive perception and has less disapproving attitudes toward cocaine use.

Vicarious learning, or modeling, occurs by observation of others' behavior, attitudes, and outcomes of the behavior and can increase the likelihood of the observer engaging in the behavior. Role models such as peers and parents affect expectancies, evaluations, and self-efficacy related to the observed behavior. Thus, even as the consequences we associate with our own behavior shape our future behavior, observing others' consequences associated with their behavior can also shape our behavior (see below for more detail).

Cognitive processes include encoding, organizing, and retrieving information, and these are postulated to regulate behavior and environmental events. An individual cognitively processes information from the environment and determines their behavioral response. Self-regulation can be defined as the ability to arrange environmental incentives and apply consequences. Thus, self-regulation need not be limited to one's ability to choose how to respond in specific situations but can also be applied to one's ability to make choices that affect the degree of exposure to specific influences. For example, if Clyde does not want to smoke marijuana and realizes that he has difficulty saying no to Paul or Mike, he can avoid Paul and Mike as a means of regulating his behavior. A related construct is self-efficacy, which can be defined as the belief that one can engage in a specific behavior and/or produce a specific desired outcome. Self-efficacy can be thought of as context-specific confidence. The extent to which Clyde believes he will be able to resist an offer to smoke marijuana is an example of self-efficacy. Self-efficacy effects and is affected by behavior. The more confidence Clyde has in his ability to resist peer influence, the more successful he will be in doing so. In turn, successful instances of resistance will increase self-efficacy to resist in the future. Furthermore, seeing a peer resist influence can boost one's self-efficacy in resisting peer influence.

Reciprocal determinism describes the associations between behavior and environmental and personal factors, each of which is affected by the other two factors. For example, peers and social environments affect subsequent smoking behaviors, and vice versa.

DIRECT AND INDIRECT INFLUENCES

Peers can contribute to addictive behaviors by way of direct (active) or indirect (passive) influences. Direct influences range from polite gestures such as "would you like a cigarette?" to overt commands such as "finish your beer." Refusal of these direct commands, especially when surrounded by peers, may result in feelings of inferiority, loss of social ease, fear of rejection and exclusion, and actual rejection or exclusion. In contrast, indirect peer influences are more subtle, and these include modeling, which is the imitation of observed behavior. For example, if someone were to observe a group of their friends using marijuana, that person would be more likely to imitate and even adopt this behavior.

Environmental influences such as social expectations contribute greatly to addictive behaviors. Research on alcohol use indicates that social expectations have a substantial influence on the association between peer and own alcohol use. An example of this phenomenon is the divergent social expectations for male and female college students with regard to drinking. In general, men are expected to drink more than women and men who drink more than other men are viewed more favorably than women who drink more than other women. In addition, although women are generally offered more drinks than their male counterparts, research indicates that women turn down offered drinks more often than men.

MODELING

Modeling – a form of indirect influence – is defined as imitation of another's behavior, presumably based on perceived contingencies of engaging in the behavior. For example, seeing another person receive social reinforcement for smoking marijuana would increase the probability that the viewer will try marijuana. During early life stages of childhood and adolescence, influential models include parents and other family members. However, during teen and adult stages of life, peers become more influential. Models with particular influence are peers such as siblings, classmates, coworkers, and friends.

Modeling research studies sometimes involve the use of confederates. A confederate is someone who participates in the experiment but has key knowledge of the nature of the study. This person is an actor who is not recognized by other participants in the study as

the hand of the researcher. A series of alcohol studies in modeling utilizing confederates have shown that participants' level of consumption of alcoholic beverages is affected by the drinking behavior of those around them. For instance, "drink matching" is the phenomenon that occurs when participants match the rate of drinking and the beverage selection of the confederate. Therefore, if the confederate chose to drink a beer rather than a cocktail, and drank half of the beer in 10 minutes, the participant would be likely to also select and consume his or her drink in a corresponding manner. To illustrate, if John were to attend a wedding during which the majority of guests were slowly sipping wine, John would likely follow their example, and similarly sip wine slowly, as opposed to doing shots or chugging beer.

Research indicates that family history and gender may play a part in susceptibility to indirect influences such as modeling. Males have been shown to better match confederates' drinking rates than females. Furthermore, having a family history of drinking problems increases the likelihood of matching to a greater degree. Thus, family history appears to moderate the susceptibility to indirect social influences on drinking. However, research shows that when the participant is with a group of sociable versus unsociable peers, the participant is likely to match the drinking rate of the group's majority, regardless of family history.

Modeling also applies to behaviors other than alcohol consumption. Indirect influences are applicable to a variety of behaviors ranging from deviant addictive behaviors like tobacco and marijuana use to socially acceptable behaviors like physical activity. For example, if someone had a group of friends who regularly exercised and approved of exercise, that person would likely be subjected to the indirect influence of modeling and at the very least, this person would probably feel some degree of pressure to adopt this behavior. Related to modeling and indirect influences, it is important to point out that individuals often have expectancies about addictive behaviors before they ever try them. For example, elementary students report alcohol expectancies before ever trying alcohol.

SOCIAL NORMS

Social norms generally refer to implicit or explicit rules regarding appropriate behavior. They provide a common framework for understanding and interpreting behavior. They represent shared values, some constant and some variable, across cultures and historical periods. Examples of social norms include wearing clothing, taking turns when talking, and not picking one's nose in public. Many of our customs and laws are designed to enforce prominent social norms, whereas others are governed more by social approval and disapproval. Social norms have a profound impact on our behavior. In some instances, such as substance use among young adults, social norms appear to be among the strongest and most consistent predictors of behavior.

Research has shown a significant link between perceived norms and addictive behaviors. For example, students who perceive that those around them support alcohol use and consume excessive alcohol are more likely to use alcohol themselves. Males tend to perceive greater support for alcohol use in their peers than females, putting males at a higher risk for succumbing to alcohol-related peer influences. Social norms have been more precisely defined upon two dimensions: perceived versus actual and descriptive versus injunctive.

Perceived versus Actual Norms

An important distinction in thinking about social norms is the difference between one's perception of the social norm and the actual social norm. This distinction is important for at least two reasons. First, we are not influenced by reality *per se*, but by our perception of reality, thus it is perceived norms, rather than actual norms that influence behavior. Second, perceptions are rarely a perfect reflection of reality. Thus, perceptions of what others approve of may or may not accurately reflect what others actually approve of. Similarly, perceptions or estimates of how many others engage in a given behavior are often inaccurate. A teen's perception that "everyone else is doing it" is almost certainly an overestimate for most behaviors in which this argument might be offered as justification for engaging in the behavior. With respect to addictive behaviors, much research has revealed systematic biases in perceived norms, particularly among, but not limited to, adolescents and young adults. When asked to estimate the level of approval and/or the prevalence of a given risk behavior such as heavy drinking, smoking, marijuana use, or other substance use, there is a strong tendency for people to overestimate, and the greater they overestimate, the more likely the perceiver is to engage in the behavior him/herself. That is, most individuals overestimate the approval of heavy drinking but heavy drinkers overestimate more than lighter drinkers or abstainers. There is also some evidence that suggests people underestimate the use of protective behaviors, such as using a designated driver.

The systematic misperception of social norms has profound implications. Behavior is heavily influenced

by perceptions of others' approval and misperceptions are often in the direction of viewing unhealthy behaviors as being more widely accepted or more prevalent. Thus, individuals may initially engage in risky behavior largely because they believe that others are engaging in the behavior or that others strongly approve of the behavior, when in fact few others may be engaging in the behavior and the level of approval may be quite low. Thus, a primary motivator of engaging in addictive behaviors is often based on faulty presumptions regarding social norms.

There are multiple possible explanations for the systematic bias between perceived and actual norms of addictive behaviors, which are not mutually exclusive. One explanation for misperception is that excessive behaviors tend to be more salient and/or memorable. Consider a party in which 100 people are present and 99 of them are either not drinking at all or are having one or two drinks and behaving appropriately. We are more likely to remember the one who is in the bathroom vomiting because he is drunk than the 99 who are behaving appropriately. When we remember the party, it is the inappropriate behavior that we remember. Relatedly, when we talk to others about the party we are less likely to talk about the 99 who were behaving appropriately but are more likely to talk about the one who was excessive. Thus, salience of excessive behavior distorts our representation of events and affects communication about behavior. Another possible explanation which has been found in some studies is a general reluctance to voice disapproval of excessive behaviors. Because silence is viewed as consent, if one person is endorsing the harmlessness of marijuana in a crowd of people, even if most people privately disapprove but are reluctant to voice their opinion, observers are likely to assume that most of those listening share approval of marijuana use. Justification and self-serving biases are other possible explanations for the bias between perceived and actual norms. Individuals who engage in excessive behavior may feel justified (i.e. less guilty) when they overestimate the prevalence of those behaviors among their peers. Similarly, individuals who do not engage in addictive behaviors, or who do so infrequently, may enhance their self-view by assuming that they are relatively unique in their abstinence or moderation. All of these explanations likely play some role in misperception, and their relative contribution may vary by behavior.

Descriptive versus Injunctive Norms

Social norms have also been described as being of two types: descriptive and injunctive. Descriptive norms, also called popular norms and behavioral norms, refer to the prevalence, quantity, and/or frequency of a given behavior. Believing that most college students engage in heavy drinking would be an example of a perceived descriptive norm. Injunctive norms refer to degree of approval or disapproval of a given behavior. They are similar to subjective norms described above but do not necessarily focus on one's own behavior or important others as a reference group. Believing that most people strongly disapprove of drinking and driving would be an example of an injunctive norm. Descriptive and injunctive norms are moderately correlated and independently associated with behavior. Recent work suggests that descriptive and injunctive norms often interact such that perceiving a behavior like smoking to be common (descriptive norm) is more strongly associated with behavior when perceived approval for the behavior (injunctive norm) is also high. Thus, an individual who believes most others disapprove of smoking will be less likely to smoke even if he or she sees many people smoking. In sum, perceived descriptive and injunctive norms are both positively associated with behavior.

RESISTANCE

Research indicates that although direct and indirect peer influences can be powerful indicators of future behavior, not everyone exposed to these influences succumbs to them. People with greater ability to resist to direct peer influence are those who are more socially confident, competent, and less socially insecure or anxious. To reiterate this, a mature, confident person with a supportive family and an established group of friends who maintains consistent, strong disapproval of drug use would be more resilient to active and passive offers of drugs even if that person were to attend a school where drug use was the norm or go to a party where drug abuse was prevalent.

Social inoculation is a construct which suggests that exposure to behaviors or attitudes heightens resilience to counterarguments. Social inoculation is thought to further increase one's resistance to the persuasive forces of social influence. For example, if a teacher wanted to inoculate his class against drug abuse, he could educate his students about the prevalence and health effects of drug use and allow the students to observe and engage in role playing the refusal of offered drugs. By doing so, the teacher would be inoculating his class against the direct and indirect influences regarding drug use, with the hope that if the students later receive drug offers from peers, they will be equipped with adequate resistance to refuse those offers.

IATROGENIC EFFECTS

Although friendships are viewed as positive relationships, these relationships can have inadvertent adverse effects due to negative peer influence. These are called iatrogenic effects. Group interventions that target behavioral problems like substance abuse have shown that iatrogenic effects can have negative outcomes for vulnerable participants. These negative outcomes can range from adoption of addictive behaviors to maladjustment as an adult. Peer aggregation among deviant individuals can promote deviancy training, which increases the likelihood of future addictive behavior. Deviancy training refers to processes by which adolescents and young adults expand their knowledge and repertoire of deviant behaviors as a function of interacting with more deviant peers. Potential mediators of deviancy regulation include normalization of less extreme behaviors, transmission of ideas, and encouragement and competition of extreme behaviors. For example, in an inpatient treatment setting, an adolescent may make friends with peers who have engaged in more extreme behaviors, and/or who present novel methods and ideas related to substance use. Moreover, exposure to deviant peers may increase accessibility of substances following treatment.

Iatrogenic effects present a challenge, but this challenge can be overcome. Although iatrogenic effects are significant, there is little evidence of these effects across research studies. Therefore, this effect is not as extensive as initially thought. Furthermore, iatrogenic effects can be constructively exploited, resulting in positive outcomes such as a decrease in problem behavior. Mixing prosocial and high-risk peers is an effective way of positively harnessing iatrogenic effects. Recent studies indicate that supportive peer involvement effectively reduces the problem behavior, especially for pairs of two people (dyads) rather than in a group setting. For example, the smoking cessation intervention would be more effective when pairing a smoker with a supportive nonsmoker, or pairing a heavy smoker with a light smoker.

SUMMARY

In summary, peer influences on addiction are central to the initiation of substance use and engagement in addictive behaviors. Peer influences seem to peak in adolescence and young adulthood. This is not to say that peers do not have a substantial influence later in life but the nature of peer relationships change considerably as emerging adults begin to transition into adult roles and responsibilities. Even when peer influences seem most evident earlier in the developmental course of life, they also seem more evident in the early developmental trajectory of addiction. Substance dependence typically includes nonsocial use and nonsocial motivations for use (e.g. tension reduction and withdrawal avoidance).

Peer influence is embedded within the larger context of social influences. A broad review of social influence theories suggests several common themes. Social influences can be explicit but they are often implicit and based on our imperfect beliefs about others' expectations of us. Moreover, others' expectations do not influence us, but rather our beliefs about others' expectations, and these are often quite inaccurate. Another common theme is that the degree of influence that others have on our behavior depends on how much we care about or identify with them. People whom we care nothing about are less likely to influence us than those with whom we closely identify. A third theme is that what is viewed as good and bad is relative to the social group. Being able to drink large quantities of alcohol is good for some and not good for others. Finally, social influences are often governed by social consequences. We observe how people respond to our own and others' behaviors and those responses accordingly increase or decrease the likelihood of future behaviors.

SEE ALSO

Cognitive Factors in Addictive Processes

Glossary

Heuristic efficient, simple rules that are learned or hard coded by evolutionary processes that explain how people make judgments, solve problems, and make decisions. Generally, heuristics are effective but can lead to biases or errors in some cases.

Iatrogenic effects inadvertent negative or adverse effects resulting from advice or treatment.

Reference group a group that acts as a frame of reference to which another group is compared when evaluating qualities such as behaviors and attitudes.

Further Reading

Abrams, D., Hogg, M.A., 1999. Social Identity and Social Cognition. Blackwell Publishing, Malden.

Blanton, H., Burkley, M., 2008. Deviance regulation theory: applications to adolescent social influence. In: Prinstein, M.J., Dodge, K.A. (Eds.), Understanding Peer Influence in Children and Adolescents. Guilford Press, New York, NY, US, pp. 94–121.

Borsari, B., Carey, K.B., 2001. Peer influences on college drinking: a review of the research. Journal of Substance Abuse 13, 391–424.

Evans, R.I., Getz, J.G., 2003. Resisting health risk behavior: the social inoculation approach and its extensions. In: Gullotta, T.P., Bloom, M. (Eds.), The Encyclopedia of Primary Prevention and Heath Promotion. Kluwer/Academic, New York, NY.

Fishbein, M., Ajzen, I., 2010. Predicting and Changing Behavior: The Reasoned Action Approach. Psychology Press, New York.

Gibbons, F.X., Gerrard, M., Blanton, H., Russell, D.W., 1998. Reasoned action and social reaction: willingness and intention as independent predictors of health risk. Journal of Personality and Social Psychology 74, 1164–1180.

Jessor, R., Donovan, J.E., Costa, F.M., 1991. Beyond Adolescence: Problem Behavior and Young Adult Development. Cambridge University Press, New York.

Lewis, M.A., Neighbors, C., Lindgren, K.P., Buckingham, K.G., and Hoang, M. (in press). Theories of Social Influence on Adolescent and Young Adult Alcohol Use. Nova Science Publishers, Inc, Hauppauge, NY.

Oetting, E.R., Beauvais, F., 1987. Peer cluster theory, socialization characteristics, and adolescent drug use: a path analysis. Journal of Counseling Psychology 34, 205–213.

Schulenberg, J., Maggs, J.L., Hurrelman, K., 1997. Health Risks and Developmental Transitions During Adolescence. Cambridge University Press, Cambridge.

Relevant Websites

http://www.addictioninfo.org – Addiction Information.

CHAPTER

34

Binge Drinking

Brian Borsari

Brown University, Providence, RI, USA

OUTLINE

CURRENT DEFINITION OF BINGE DRINKING

The term binge drinking has appeared in print for many years, although a precise definition did not exist. In 2004, The National Institute of Alcohol Abuse and Alcoholism (NIAAA) attempted to reduce the ambiguity of previous definitions used in popular society by proposing a definition of binge drinking. On February 5, 2004, the NIAAA National Advisory Council approved the following definition for binge drinking, which had been put forth by a special task force that had been assigned the responsibility of defining binge drinking and differentiating it from other patterns of alcohol use. This task force met in a 2-day workshop held in November 2003, in which experts in neurology, psychology, and physiology discussed how to best define and measure binge drinking. The following definition resulted from these efforts:

> *A binge is a pattern of drinking alcohol that brings blood alcohol concentration to 0.08 g% or above. For the typical adult, this pattern corresponds to consuming five or more drinks (male), or four or more drinks (female) in about 2 h. Binge drinking is clearly dangerous for the drinker and for society.*

This definition was accompanied by the following clarifications: (1) a "drink" refers to half an ounce of alcohol (e.g. one 12oz. beer, one 5oz. glass of wine, or one 1.5oz. shot of distilled spirits); (2) binge drinking is different from risky drinking (reaching a peak BAC of 0.05 g% to 0.08 g%) and a bender (2 or more days of sustained heavy drinking); (3) for some individuals (e.g. older people or people taking other drugs or certain medications), the number of drinks needed to reach a binge-level BAC is lower than the "typical adult"; (4) people with risk factors for the development of alcoholism have increased risk with any level of consumption; (5) for pregnant women, any drinking presents risk to the fetus; and (6) drinking by persons under the age of 21 is illegal.

Since this announcement was published in the NIAAA newsletter, this has been the standard definition for binge drinking in the United States, and it has been adopted in other countries as well, especially in the context of adolescent and college drinking research.

Principles of Addiction, Volume 1
DOI: http://dx.doi.org/10.1016/B978-0-12-398336-7.00034-6

HISTORY OF THE TERM

Multiple definitions of binge drinking exist and the meaning of the term varies according to the source and the era. For example, the Oxford English Dictionary defines binge drinking as the consumption of an excessive amount of alcohol in a short time, Collins English Dictionary defines it as the practice of drinking excessive amounts of alcohol regularly, and Dictionary.com describes the behavior as the consumption of dangerously large quantities of alcoholic beverages in one sitting. All three of these definitions indicate that binge drinking occurs in one occasion, but these definitions fail to provide a concrete definition of what excessive amounts, regularly, and dangerously large.

Until approximately 20 years ago, a binge was often thought of to be an intense, multiday or week-long period of drinking that often was done in a solitary fashion. The purpose of binge drinking was to become intoxicated, and a loss of control was a component of such a binge drinking episode. The binge was often accompanied by significant consequences such as blackouts, injuries and often ended in incarceration or inpatient treatment. Although a binge was occasionally referred to as a single event, the amounts were usually extreme (e.g. 12 or more drinks). Furthermore, binge drinking episodes, or binges, would often be separated by extended periods of sobriety.

The conceptualization of a binge as a discrete episode of extreme drinking is consistent with Jellinek's epsilon alcoholism, or periodic alcoholism, marked by the individual's primary drinking style being binge drinking. This form of alcoholism was relatively rare, however, and Jellinek suggested that it might have been a symptom of another disorder. Further confusion was caused by the fact that binge drinking episodes could occur in the context of other types of alcoholism: alpha (psychological dependence on alcohol), beta (alcohol-related consequences such as cirrhosis occur in the absence of psychological or physiological dependence, proposed to occur in cultures in which heavy drinking is accepted), gamma (progression from psychological to physiological dependence with loss of control), and delta (psychological and physiological dependence without lack of control). Further confusing the issue was the term bender, or bender alcoholics. Benders were popularly conceptualized as a drinking bouts of 48 h or more, with a failure to meet work, family, or other obligations.

Therefore, for many years, the terms binge drinking and bender were used colloquially and in research in the United States. Through the 1980s, the term binge drinking was used to describe a single incident of drinking five or more drinks per occasion in research, but there was still no formal consensus on the definition of the term. However, the term became popular in the media and in research in relation to college students started in the early to mid-1990s. Specifically, Dr Henry Wechsler of the Harvard School of Public Health had started to conduct large-scale surveys of college student alcohol use – the College Alcohol Survey. Dr Wechsler had been studying heavy drinking in college students for over a decade and had first used the term in a 1991 report. The College Alcohol Survey, funded by grants from the Robert Wood Johnson Foundation, permitted a large-scale examination of this phenomenon. The CAS surveyed over 50,000 students at 140 colleges and universities in the United States. The first of these surveys was conducted in 1993, and subsequent surveys were conducted in 1997, 1999, and 2001. Originally defined as having five drinks per occasion, the definition was later amended to five drinks on one occasion for men and four drinks on one occasion for women. Through the many research articles published on the topic from this data, and subsequent research, "binge drinking" was conceptualized as the threshold for which alcohol-related harms became significantly more likely for the drinker, as well as those who share the college environment with the drinker. These "secondary effects" were also a source of concern for school administrators, and a commonly cited reason to increase prevention and intervention efforts on college campuses.

As the term binge drinking appeared with greater frequency in research and the popular press, often coinciding with the release of results of another CAS survey, there was considerable debate about the definition of binge drinking. One benefit of the term binge drinking is the potential to quickly provide descriptive information about a drinking episode. However, opponents of the term binge drinking definition argued that the use of the term binge drinking was not always synonymous with intoxication prior to NIAAA's definition in 2004 and that this threshold was not an appropriate marker of excessive drinking for everyone. In addition, the use of the term binge drinking in the media was believed to sensationalize alcohol use to the point where it might contribute to inaccurate normative perceptions about the frequency of heavy intoxication. Therefore, there was (and continues to be) debate regarding the definition of binge drinking that is markedly different from the historical definition. Synonyms of the term binge drinking were introduced to differentiate the historical and the current definition for binge drinking, and these synonyms included heavy episodic drinking, risky drinking episode, harmful drinking, dangerous drinking, excessive drinking, and problem drinking.

Concurrent with the focus on binge drinking in the United States, there was a large-scale refinement of the phrase binge drinking as researchers and policy makers

around the world attempted to define what exactly constituted binge drinking. As a result, efforts were made to shift the focus to single days, or single drinking occasions, and also to determine how many drinks needed to be consumed to constitute a binge. However, to date, there is no worldwide consensus on the amount of alcohol that needs to be consumed to qualify as a binge drinking episode. Indeed, global cutoffs include a half bottle of spirits or two bottles of wine on one occasion (Sweden); double the daily recommended amount of alcohol – about 2–3 drinks (or 1 or 2 for women) (England); six bottles of beer (Finland); and 6 or more units of alcohol (women) and 10 or more units of alcohol for men (United Kingdom).

In addition to a lack of consensus regarding the amount of alcohol required to constitute a binge, there was still omitted information that directly influences the degree of intoxication, related risks, and cultural issues. First, a time frame in which the alcohol was consumed was not specified. For example, the threshold of drinks could be attained in 5 min or 5 h, with clear implications on the degree of intoxication. Later, the phrase in one sitting was often included but did not completely address the issue. Second, the body weight of the individual was not taken into account. As a result, the degree of intoxication from a binge drinking episode was difficult to discern. Third, there have been no attempts to account for the effect of previous or concurrent food consumption on intoxication. Finally, the motivation for binge drinking can vary – in some situations getting intoxicated is the desired outcome (e.g. in college students) and in others situations binge drinking may be unintentional or related to mental illness. Taken together, these issues make it difficult to precisely interpret research assessing binge drinking episodes. That said, researchers around the world continue to examine different drink cutoffs and definitions for binge drinking, seeking to find the definition that best links levels of intoxication with consequences.

The term binge drinker has often been used, but the precise definition of what makes someone a binge drinker has remained vague. Therefore, there have been various attempts to define a binge drinker by the frequency of binge drinking episodes he/she engages in over a certain time period. Time frames used to capture binge drinking have ranged from the past week to the past year. Overall, a 6-month time frame has been determined by many to be most informative in linking binge drinking to alcohol-related consequences. Dr Wechsler and colleagues further differentiated between binge drinkers (one or two binge drinking episodes in the past 2 weeks) and frequent binge drinkers (three or more binge episodes in the past two weeks). However, this definition has not been universally accepted.

Because of the lack of a stable definition of the term binge drinking, care must be taken to clearly define binge drinking when interpreting or disseminating research findings, especially from different studies. However, as the definition of binge drinking has been clarified over the past 30 years, research has implemented the item into interviews and surveys with a variety of populations. It is possible to identify some trends in the literature on binge drinking (defined in most studies as five or more drinks on one occasion for men and women).

BINGE DRINKING IN ADOLESCENTS

Binge drinking typically peaks in adolescence and declines. Examination of large-scale surveys assessing binge drinking indicated an overall decrease in binge drinking in adolescents and young adults over the past 30 years. However, there was one exception: the rates of binge drinking in college students remained stable. One of the common findings in the research with college students was that often the binge drinking observed on campus was a continuation of binge drinking that had occurred in high school. Some other trends were apparent, as well: women seem to be engaging in binge drinking more, and binge drinking decreased among minority men (but not women).

Research with adolescents consistently indicates that binge drinking is related to injuries, violence, driving while intoxicated, unsafe sexual practices, and death. In addition, binge drinking is the most common style of drinking among adolescents, accounting for over 90% of the alcohol consumed by high school students. This is of concern, as establishing this pattern of alcohol use early in life may lead to continued binge drinking throughout the lifespan, as well as increase the risk for developing alcohol dependence. The association of binge drinking with the concurrent use of other substances, such as tobacco, has also been repeatedly observed.

Recent research has raised worries about the long-term effects of binge drinking on the developing brain; specifically, adolescent brains may be more susceptible to damage from binge drinking than those of adults. Specifically, research has indicated that binge drinking (2–3 times per month, at least 4–5 drinks per occasion) can negatively influence adolescents' performance on thinking and memory tests. Furthermore, it seems that binge drinking may influence the white matter in the brain, making communication among brain cells less efficient. Other areas of the brain, such as the hippocampus, may also be compromised by binge drinking. The long-term effects of this impairment, and whether

it can be partially or totally reversed, have yet to be determined.

BINGE DRINKING IN YOUNG ADULTS

Binge drinking is most prevalent in the 18- to 20-year-olds, and this behavior appears to be especially prevalent in college students. Large-scale surveys indicate that approximately 68% of all college students drank alcohol in the past month, and 40% of college students engage in binge drinking. Recent studies indicate that alcohol use by college students has increased in the late 1990s and early 2000; for example, cross-sectional data showed a 16% increase in the proportion of students who engage in frequent heavy episodic drinking ($\geq$3 in the past two weeks) between 1993 and 2001. Heavy episodic drinking is related to numerous alcohol-related consequences in college students, including academic difficulties, property damage, risky sexual activity, blackouts, alcohol poisoning, and death.

In 2009 and 2010, attention was placed on drinking behaviors labeled extreme binge drinking or industrial-strength bingeing. This practice involves drinking very large quantities of alcohol (10 or more drinks) over a short time (less than an hour), in order to achieve high levels of drunkenness. This practice has yet to receive any formal research attention, however, so its prevalence and associated risks are not known. However, the emergence of this practice indicates that binge drinking and its conceptualization will continue to evolve.

BINGE DRINKING IN ADULTS

The consequences associated with binge drinking in adolescents and college students extend into adulthood. Binge drinking is common in adults, and more than 50% of the alcohol consumed by adults is consumed in the context of binge drinking episodes. In addition, men tend to binge drink more than women, accounting for as many as 81% of adult binge drinking episodes.

BINGE DRINKING IN THE ELDERLY

Binge drinking rates are lowest in adults aged 65 and older, although the majority of the risks associated with younger adults are still present in this population (e.g. drinking and driving, injuries). There may be a unique consequence in this population; however, preliminary evidence suggests that binge drinking may be associated with the onset of dementia. Consistent with younger adults, men are more likely to binge drink, and Caucasian men are most likely to binge drink, followed by Hispanic and African American men. This trend is reversed in women, with African American women exhibiting more binge drinking than Caucasian or Hispanic women. For both men and women, however, social isolation and the lack of an observed connection between psychological distress and binge drinking may make detection and intervention difficult by family, friends, and caregivers.

GENDER DIFFERENCES IN BINGE DRINKING

Across all age groups, men consistently report higher levels of binge drinking than women, even with the gender-specific definition. Binge drinking in pregnant women can cause significant danger to the fetus such as fetal alcohol spectrum disorders (a risk also relevant to adolescents and young adults).

ETHNIC DIFFERENCES IN BINGE DRINKING

Surveys across all age groups indicate, with little fluctuation, that Caucasians binge drink more than any other racial group. Hispanics report the second-highest rates of binge drinking, followed by African Americans. Asians exhibit the lowest rates of binge drinking, although research has indicated that among Asian subgroups, Koreans and Filipinos report the highest rates of binge drinking.

CULTURAL DIFFERENCES IN BINGE DRINKING

Alcohol has a role in each culture, and the practice of binge drinking has different levels of acceptance depending on the culture in which it occurs. For example, in China, binge drinking rates are generally quite low (e.g. 4% of the population). However, at banquets, official functions, or business meetings, many rounds of spirits are often consumed with the toast of gambei (dry glass). The refusal to drink can be interpreted as an insult, leading to social or business repercussions. This pattern of consumption can lead to extreme intoxication, and the deaths of several government officials led the Chinese Communist Party to develop a formal code of conduct for officials as well as ban lavish weddings and funerals in 2010. However, binge drinking is also becoming more prevalent in adolescents and young adults in China.

A similar pattern of binge drinking is also found in Japan, where business meetings are often accompanied with heavy drinking. A common saying iki-iki-iki (drink-up, drink-up) is a verbal cue to have individuals finish their drink in one gulp. Alcohol is highly prevalent in Japanese culture, and binge drinking is seen as part of business. Refusing offers to drink with one's boss may lead to not being promoted. In the larger culture, becoming intoxicated is seen as a way to relax, speak one's mind, or get to know others. Public drunkenness is also more highly tolerated in Japan than other countries. Certain times of the year are also associated with heavy drinking, particularly nenmatsu (the year's end) and nenshhi (the year's beginning). Japanese alcohol consumption has decreased since peaking in the late 1990s, partially attributed to an aging population and a change in societal norms toward binge drinking.

In Europe, eastern and northern Europeans seem to engage in, and be more accepting of, binge drinking and drunkenness. For example, drinking to get drunk is commonly practiced in the United Kingdom, and in Russia binge drinking (zapoy) usually consists of two days of continuous drunkenness. In contrast, southern Europe – France, Italy, and the Mediterranean counterparts – are less tolerant of binge drinking and public drunkenness.

PREVENTION AND INTERVENTION EFFORTS

As the definition of binge drinking has evolved, as have the efforts to reduce or stop this behavior in individuals of all ages. These efforts include educating healthcare workers about detecting binge drinking and informing patients about its associated risks, raising the drinking age, limiting access to alcohol, and the increased enforcement of drinking and driving laws. As the concern about binge drinking on college campuses increased throughout the 1990s in the United States, social norms marketing (SNM) campaigns became a widespread approach to address binge drinking in college students. This approach proposes that heavy drinking in students is influenced by their misperception of other students' drinking. Therefore, using mass media to educate students about actual rates of binge drinking (if lower than the perceived rates) will therefore reduce on-campus alcohol use, as the students would feel less pressure to conform to the perceived norm. In this way, SNM campaigns attempted to change the culture of heavy drinking on college campuses. Results for this approach have been mixed. In clinical settings, individual and group brief motivational interventions (BMIs) typically provided to individuals currently engaging in binge drinking and experiencing problems, provide skills to resist offers of alcohol from peers, alternate behaviors to drinking, and educate the individual about the risks associated with binge drinking. These interventions have been administered in a wide variety of contexts, formats, and settings, and research indicates a consistent small to moderate effects on the frequency of binge drinking and related consequences. One area of particular innovation is web-based interventions, which have demonstrated efficacy with college students and adults and have fostered continued efforts to refine approaches to screening, assessment, and intervention through the internet.

SEE ALSO

Adolescent Substance Use: Symptoms and Course, Symptoms and Course: Alcohol Use Disorder in Adulthood, Symptoms and Course: Older Age and Substance Abuse, Substance Use and Mental Health Issues on the College Campus

Glossary

BMIs brief motivational interventions.

Binge this term has had a variety of interpretations, but commonly conceptualized to be an intense episode of drinking, the purpose of which was to become intoxicated.

Binge drinking this term has had various definitions over the years, but was formally defined in 2004 by the National Institute of Alcohol Abuse and Alcoholism. A pattern of drinking alcohol that brings blood alcohol concentration to 0.08 g% or above. For the typical adult, this pattern corresponds to consuming five or more drinks (male), or four or more drinks (female) in about 2 h. Binge drinking is clearly dangerous for the drinker and for society. This definition has not been accepted universally by clinician and researchers, however.

Binge drinker as defined by Dr Wechsler, binge drinkers are individuals who have had five or more drinks in a row (four drinks for females) at least once in a 2-week period. This definition of a binge drinker has been further refined to take into account the frequency of binge drinking. Specifically, an occasional binge drinker is defined as an individual who has had five or more drinks in a row (four drinks for females) once or twice in a 2-week period, and a frequent binge drinker is an individual who has had five or more drinks in a row (four drinks for females) three or more times in a 2-week period.

Bender a two or more days episode of heavy drinking, often accompanied by blackouts or loss of consciousness.

Heavy episodic drinking term used to describe a drinking incident that leads to drunkenness, and the amount consumed varies among definitions.

NIAAA National Institute of Alcohol Abuse and Alcoholism.

Risky alcohol use defined in 2004 by the National Institute of Alcohol Abuse and Alcoholism as a drinking event in which an individual's peak blood alcohol level of 0.05 g% is obtained.

SNM social norms marketing campaigns.

Further Reading

Carey, K.B., 2001. Understanding binge drinking: introduction to the special issue. Psychology of Addictive Behaviors 15, 283–286.

Courtney, K.E., Polich, J., 2009. Binge drinking in young adults: data, definitions, and determinants. Psychological Bulletin 135, 142–156.

Herring, R., Berridge, V., Thom, B., 2008. Binge drinking: an exploration of a confused concept. Journal of Epidemiology and Community Health 62, 476–479.

Hingson, R.W., Edwards, E.M., Heeren, T., Rosenbloom, D., 2009. Age of drinking onset and injuries, motor vehicle crashes, and physical fights after drinking and when not drinking. Alcoholism, Clinical and Experimental Research 33, 783–790.

Jackson, K.M., Sher, K.J., Gotham, H.J., Wood, P.K., 2001. Transitioning into and out of large-effect drinking in young adulthood. Journal of Abnormal Psychology 110, 378.

Lange, J.E., Clapp, J.D., Turrisi, R., Reavy, R., Jaccard, J., Johnson, M.B., et al., 2002. College binge drinking: what is it? Who does it? Alcoholism, Clinical and Experimental Research 26, 723–730.

National Institute on Alcohol Abuse and Alcoholism, 2004. NIAAA council approves definition of binge drinking. NIAAA Newsletter 3, 3.

Parra, G.R., Krull, J.L., Sher, K.J., Jackson, K.M., 2007. Frequency of heavy drinking and perceived peer alcohol involvement: comparison of influence and selection mechanisms from a developmental perspective. Addictive Behaviors 32, 2211–2225.

Sher, K.J., Bartholow, B.D., Nanda, S., 2001. Short- and long-term effects of fraternity and sorority membership on heavy drinking: a social norms perspective. Psychology of Addictive Behaviors 15, 42–51.

Sher, K.J., Rutledge, P.C., 2007. Heavy drinking across the transition to college: predicting first-semester heavy drinking from precollege variables. Addictive Behaviors 32, 819–835.

Wechsler, H., Davenport, A., Dowdall, G.W., Moeykens, B., Castillo, S., 1994. Health and behavioral consequences of binge drinking in college: a national survey of students at 140 campuses. Journal of the American Medical Association 272, 1672–1677.

Wechsler, H., Isaac, N., 1992. "Binge" drinkers in Massachusetts colleges: prevalence, drinking style, time trends, and associated problems. Journal of the American Medical Association 267, 2929–2931.

Wechsler, H., Lee, J.E., Kuo, M., Seibring, M., Nelson, T.F., Lee, H., 2002. Trends in college binge drinking during a period of increased prevention efforts. Findings from 4 Harvard School of Public Health College Alcohol Study surveys: 1993–2001. Journal of American College Health 50, 203–217.

Wechsler, H., Nelson, T.F., 2001. Binge drinking and the American college students: what's five drinks? Psychology of Addictive Behaviors 15, 287–291.

Wechsler, H., Wuethrich, B., 2002. Dying to Drink: Confronting Binge Drinking on College Campuses. Rodale, Emmaus, PA.

Relevant Websites

http://kidshealth.org/teen/drug_alcohol/alcohol/binge_drink.html – Overview of binge drinking in adolescents.

http://www2.potsdam.edu/hansondj/index.html – Alcohol: problems and solutions site, maintained by Dr David Hanson.

http://www.hsph.harvard.edu/cas/About/index.html – Dr Wechsler's College Alcohol Survey website.

http://www.niaaa.nih.gov/Pages/default.aspx – The National Institute of Alcohol Abuse and Alcoholism website.

CHAPTER

35

Alcohol's Effects on Sexual Arousal and Sexual Functioning

William H. George, Amanda K. Gilmore

University of Washington, Seattle, WA, USA

OUTLINE

OVERVIEW

Alcohol and sexuality are closely linked in art, history, popular culture, anecdotal experience, and clinical lore. The sheer ubiquity and durability of these linkages suggest that something about drinking alcohol systematically co-occurs with or systematically influences something about sexual responding. Scientists have investigated these possibilities using a variety of research methods including questionnaire surveys, personal interviews, and laboratory experiments. Research findings indicate that alcohol is reliably associated with and can exert a causal impact on a variety of sexual processes and outcomes, including – for example – the degree to which one perceives situations and events as sexual in nature and one's willingness to engage in risky sexual decisions, such as unprotected intercourse with a new partner. However, there is a more basic question that potentially underpins these connections from alcohol to sexual processes and outcomes: How does acute or chronic drinking affect sexual functioning? That is, does alcohol intoxication affect sexual desire, arousal, orgasm, and sexual pain; if so, how; and do these potential effects contribute to clinically diagnosable levels of distress? These questions are the focus of the current chapter.

Alcohol is related to variability in sexual functioning in both men and women. The relationship between alcohol and sexual functioning is so evident in both community and clinical samples that the disorder *Substance-Induced Sexual Dysfunction* is in the Diagnostic and Statistical Manual of Mental Disorders, fourth edition text revision. Survey research indicates that women report more problems with sexual functioning than men. However, alcohol use rates are typically higher among men than women. In studies with men,

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00035-8

heavy drinking and alcohol abuse have been found to be associated with more sexual dysfunction and less sexual satisfaction, yet these effects are thought to diminish with sobriety. In studies with women, heavy drinking and alcohol abuse have been associated with sexual dysfunction. However, these studies only imply an association and do not indicate causality. Therefore, it is important to examine both survey and experimental research when discussing alcohol's impact on sexual functioning. For example, it is possible that an individual may use alcohol to cope with a sexual dysfunction because of popularly held beliefs that alcohol enhances sexuality. However, because alcohol can have the opposite effect on sexual functioning, alcohol use to cope with sexual dysfunction may actually decrease sexual functioning. This chapter will discuss the relationship between alcohol and sexual functioning. Because sexual functioning differs based on gender, we will examine research regarding men and women separately.

DEFINITIONS

It is worthwhile to specify several definitions at the outset. *Sexual functioning* can be defined as a normative sexual response including sexual desire, sexual excitement and arousal, orgasm, and resolution. For the purposes of this chapter, we will focus on sexual function in relation to sexual desire, sexual arousal, orgasm, and sexual pain. *Sexual dysfunctions* are defined as disturbances in the normal sexual response that also cause distress and difficulties interpersonally. In contemplating alcohol's effects on sexual functioning, it is important to distinguish between acute alcohol consumption versus chronic consumption. Acute consumption refers to the amount of alcohol consumed in a singular index drinking episode and the resultant effects of alcohol on sexual responding in the context of that index episode. Acute alcohol consumption is operationalized differently in survey and interview studies compared to experimental studies. In survey and interview studies, acute alcohol consumption refers to retrospective reports of the amount consumed in an episode and the sexual events occurring during or after the drinking episode. Alternatively, in experiments, acute consumption is operationalized by having research participants come into a controlled laboratory setting and ingest beverages containing predetermined dosages of alcohol designed to result in precisely targeted blood alcohol concentrations. In contrast with survey and interview studies, which yield primarily correlational findings, such experiments on acute alcohol effects have been uniquely informative about the existence and nature of alcohol's causal impact on sexual functioning. Chronic alcohol consumption refers to the amount of alcohol consumed habitually over time and is operationalized based on research participants' retrospective reporting. These studies, while less informative about causality, have been uniquely informative about alcohol's associations with sexual functioning in real life – as opposed to strictly laboratory settings. Within the domain of chronic alcohol consumption, it is important to distinguish two levels of problematic drinking: *Alcohol abuse* and *alcohol dependence,* which are diagnosable clinical conditions. Alcohol abuse refers to chronic use that results in a number of social, emotional, or physical problems. Alcohol dependence, the more severe condition of the two, refers to chronic use that results in numerous problems and is often accompanied by physical dependence and alcohol tolerance symptoms.

ALCOHOL AND SEXUAL DESIRE

Women

Approximately 30% of women in the United States are affected by problems with sexual desire to a clinical level. The current research on women's alcohol use and sexual desire is limited, and the few studies that have examined this relationship have had contradictory findings and consist predominately of descriptive and survey research. The majority of research has found that alcohol is related to increased sexual desire in women. However, studies have also found the opposite effect as well as null effects on sexual desire. Studies supporting the finding that alcohol is related to increased sexual desire have included both alcohol dependent samples and community samples without alcohol problems. One study specifically examined red wine intake in Italian women and found that women who reported moderate daily red wine intake also reported higher levels of sexual desire compared to women who did not drink alcohol. Studies finding the opposite effect, that alcohol inhibits sexual desire, have primarily been women with alcohol problems. One study found that although alcohol-dependent women were more likely to believe that alcohol enhanced sex compared to women who were not alcohol dependent, they were also more likely to have a lack of interest in sex. Finally, one study using an epidemiological sample found that inhibited sexual desire was not associated with alcohol use for men or women. No study that we are aware of has examined alcohol and women's sexual desire in an experimental paradigm.

We currently do not know why there are contradictory findings of alcohol's effect on women's sexual desire. One possible reason may be sex-related alcohol expectancies (beliefs women hold about the effects of

alcohol on sex). It is a commonly held belief that alcohol increases sexual desire and enhances sex, thus a woman might experience more sexual desire when intoxicated if she holds that belief. However, experiments indicate that physiologically, alcohol typically has an inhibiting effect on the sexual functioning cycle of arousal and orgasm. It is possible that studies finding that alcohol increases sexual desire are capturing an alcohol expectancy effect. However, further research using experimental designs would be useful to gain a better understanding of these contradictory findings.

Men

Approximately 15% of men are affected by problems with sexual desire to a clinical level. While there has been extensive research on the relationship between alcohol and sexual arousal, little is known scientifically about the relationship between alcohol and men's sexual desire, except that chronic heavy drinking can reduce sexual interest and desire. Anecdotal and clinical indications are that acute alcohol intoxication has a disinhibitory effect, increasing men's sexual interest and desire. Cross-sectional survey research indicates that men – and women – generally expect that alcohol disinhibits and enhances sexual feelings and behaviors. Men experience this as true for self and for others, as true in the past, and as likely to be true in the future. Only a handful of relevant experimental studies have been conducted. Although sexual desire was not assessed *per se*, experimental findings have indicated that acute alcohol consumption variables – including placebos, low to moderate dosages, explicit situational alcohol cues, and implicit alcohol priming cues – result in heterosexual men reporting and exhibiting greater interest in erotic material, reporting greater arousal, and rating women as more sexually attractive. Thus, at present, it seems clear that men expect that acute alcohol increases sexual desire and when exposed to alcohol acutely they respond more strongly on indices related to sexual desire. More experimental research is needed, which directly measures sexual desire independent of arousal and evaluates alcohol's direct effects on desire.

ALCOHOL AND SEXUAL AROUSAL

Women

Approximately 10–30% of women in the United States report problems with sexual arousal to a clinical level. Research on alcohol and sexual arousal in women has included both survey and experimental studies. Similar to research on women's sexual desire, the few survey studies that have examined sexual arousal and alcohol use in women have had mixed findings. One study comparing alcohol-dependent women in treatment to community women without alcohol problems found that alcohol-dependent women were more likely to have a lack of arousal compared to the women with no alcohol problems. In a daily diary study, alcohol was found to have no effect on arousal. However, in another diary study, it was found that alcohol use was related to sexual arousal during the post- and inter-menstrual phases of the menstrual cycle. Although the survey research examining the relationship between sexual arousal and alcohol use in women has been contradictory, two studies suggest that sexual arousal increases with sobriety in women. This suggests that it is possible for women to recover from an alcohol-induced diminution of sexual arousal by becoming sober.

Experimental research has also focused on alcohol's effects on sexual arousal. These studies have focused both on self-reported sexual arousal and genital arousal. As of 2011, a total of 15 experiments have examined self-reported sexual arousal in the context of acute alcohol intoxication. The majority, specifically 13 out of 15, of these studies found that alcohol intoxication increased self-reported arousal in women. The effect of alcohol has been more pronounced with higher dosages, but the effect was present at a variety of doses. In addition, the effect has been shown with a variety of different stimuli including erotic films, vignettes, and slides. As discussed with respect to sexual desire, one explanation for alcohol's effect on increasing self-reported sexual arousal in women is alcohol expectancy theory. This theory suggests a self-fulfilling prophecy whereby individuals who hold the belief that alcohol increases sexual arousal are more likely to perceive that alcohol enhances sexual arousal in alcohol-involved sexual situations than those who do not hold that belief. A second explanation for increased self-reported sexual arousal may be alcohol myopia theory, which asserts that alcohol-induced cognitive impairment constricts the range of cues one is capable of attending to in a situation. Alcohol myopia theory would therefore posit that intoxicated individuals may focus more intently on salient sexual cues such as erotic stimuli than sober individuals and consequently experience themselves as more aroused. A third possible explanation for the increase in self-reported sexual arousal when intoxicated with low levels of alcohol may be an increase in testosterone levels. Acute alcohol research conducted with Finnish women has found that alcohol's effects on self-reported sexual arousal occur in specific phases of the menstrual cycle. These findings together suggest that perhaps alcohol's effects on self-reported sexual arousal may be related to testosterone levels. These potential explanations for how low to moderate dosages of alcohol can

enhance women's self-reported sexual arousal warrant further research.

Nevertheless, regardless of the explanation, it seems clear that low to moderate dosages of alcohol result in women feeling and reporting more sexual arousal than if they were sober. Paradoxically, it seems that this effect can occur concurrently while alcohol is actually attenuating genital arousal. That is, in some experiments, women reported post-drinking increases in sexual arousal; yet, they also exhibited alcohol-induced decreases in physiologically measured genital arousal. This paradox is not surprising given the low agreement generally between women's self-report and genital measures across all sexual arousal studies (not just alcohol studies). In a review of studies examining the correspondence between self-report versus genital arousal, self-report and genital measures correlated substantially lower for women ($r = .26$) than for men ($r = .66$). Therefore, although paradoxical, it is not contradictory empirically that alcohol can have opposite effects on women's self-reported versus genital arousal.

Experiments evaluating alcohol's effects on genital arousal typically use a vaginal photoplethysmograph as measurement. This device uses light to detect the amount of blood flow in the vaginal walls, an index of lubrication and arousal. Of 10 such experiments that have been reported to date, 6 found that alcohol attenuated women's genital arousal and 4 found no evidence of attenuation. In two recent illustrative experiments with women between the ages of 21 and 35 years, the effects of two alcohol dosages (.08 and .10 mg%) were evaluated in the context of instructions to control arousal by either maximizing or suppressing it while viewing an erotic film. At the lower dosage, which is widely recognized as substantially intoxicating in that it is the legal criterion for drunk driving and it is widely associated with generalized impairment, alcohol did not attenuate genital arousal. Furthermore, women exhibited greater genital arousal under conditions of being instructed to maximize arousal than to suppress arousal. These findings were interesting in demonstrating that, despite being intoxicated and despite the generically low correspondence between subjective and genital arousal, women were able to exert control over their genital arousal. However, at the high dosage, alcohol did attenuate women's genital arousal and no arousal control was evident. Thus, it seems clear that alcohol can attenuate women's genital arousal and this effect is most evident at very high dosages. As noted earlier, the mechanisms responsible for this attenuation effect have not been clearly identified. It has been presumed that the mechanisms are strictly physiological. However, this has not been established and psychological mechanisms have not been ruled out.

Looking across both the self-report and genital studies, it appears reasonable to contend that until the blood alcohol content exceeds .08 mg%, women are likely to exhibit an arousal response pattern characterized by enhanced self-reported sexual arousal, no alcohol-induced diminution in genital arousal, and an effective capability to control their arousal response volitionally. These considerations suggest that women's post-drinking sexual responses are subject to considerable variability and are not dictated by physiological imperatives or limits. However, this variability apparently decreases at higher dosages where it appears that intoxication effects may become less conducive to positive sexual responding and experience. Potentially, physiology related to high alcohol levels may overwhelm the sexuality-related response systems.

A few studies have found that alcohol is indirectly associated with sexual risk through self-reported sexual arousal, suggesting that alcohol's effects on sexual arousal can have an impact on other aspects of women's sexual health. In addition, more recent studies have found that a history of child sexual abuse or adult sexual assault also has an impact on sexual arousal. Because a history of sexual assault is also associated with increased alcohol use, these two factors may co-occur and have an impact on women's sexual arousal.

Men

Considerable research has been conducted examining the relationship between acute alcohol use and men's sexual arousal. Appraising this research necessitates highlighting key methodological distinctions critical to delineating and understanding the association between alcohol use and men's arousal. These include acute versus chronic alcohol consumption effects and alcohol's physiological effects versus its learned – expectancy – effects.

Regarding chronic alcohol consumption, it has become well established that regular consumption of high amounts of alcohol has been consistently linked with impaired erectile functioning and associated sexual dysfunctions in men. Samples of alcoholic men exhibit greater prevalence of erectile disorder than nonalcoholic samples. Similarly, among samples of men seeking treatment for erectile dysfunction, moderate to high alcohol use has been associated with worse sexual function. Perhaps because of such findings obtained with alcoholic and sexually dysfunctional men, it has been presumed that chronic consumption of low and moderate amounts of alcohol also had deleterious effects on erectile functioning for all men. This presumption was in spite of survey findings and ample anecdotal evidence indicating that modest alcohol consumption was associated with enhanced sexual functioning.

However, two recent review papers evaluating numerous population-based studies examining the effects of chronic alcohol consumption on erectile functioning have indicated that these survey and anecdotal findings have merit. Indeed, in both review papers, authors concluded that consuming low to moderate alcohol amounts actually offered some protection against experiencing erectile dysfunction. This finding is more in keeping with survey and anecdotal evidence and suggests that chronic drinking becomes a problem for erectile responding only when consumption amounts are high.

Thus, recent evidence has ushered a shift in our understanding of chronic alcohol consumption and erectile functioning, from appraising all consumption as deleterious to recognizing nondeleterious effects – and perhaps even salutary effects – for low to moderate consumption. There has been a similar shift in our understanding of the effects of acute alcohol consumption on erectile functioning. While it once seemed conclusive that moderate to high doses of alcohol administered acutely attenuated men's erectile responding, recent research has illuminated more limiting and qualifying parameters on this so-called attenuation effect.

Acute Alcohol Attenuation of Erectile Responding: Then and Now

Approximately two dozen highly controlled laboratory experiments have been conducted examining the effects of acute alcohol consumption on erectile responding. Evidence of an attenuation effect – alcohol reduces erection magnitude – was initially established based on findings from three of four experiments published in 1976. In these pioneering experiments, young (mostly college) men were administered alcoholic or nonalcoholic drinks and presented with explicit sexual stimuli while erectile responding was assessed using a strain gauge to measure penile circumference – an indicator of erection magnitude. The findings revealed that, except at very low dosages, alcohol attenuated erectile response. Intoxicated men had shown less penile tumescence and presumably less erection magnitude than sober men. Another early study showed that alcohol also increased the amount of time it took to achieve orgasm. After this pioneering work, alcohol-induced attenuation of erectile responding became axiomatic. Three subsequent reports corroborated this finding.

Two aspects of this attenuation effect were especially noteworthy. First, it cohered with what were understood, at the time, to be the attenuating effects of chronic alcohol consumption. Thus, the scientific conclusion seemed cohesive and unambiguous: Both chronic drinking and acute (except for very low dosages) drinking undermine erectile performance. This conclusion, of course, rendered as counterfactual the longstanding lore about alcohol's aphrodisiac qualities disinhibiting men's sexual responding. Second, in light of the emerging work on alcohol expectancy effects in the 1970s, the attenuation effect seemed to contribute to completing a cohesive narrative capable of resolving the old Shakespearean paradox about alcohol and sex: "Lechery, sir, it [drink] provokes, and unprovokes: it provokes the desire, but it takes away the performance" (Macbeth, 2.3.32). Specifically, it seemed that alcohol's salutary – "it provokes" – effects were attributable to its psychological expectancy properties, whereas it's deleterious – "it unprovokes" – effects were attributable to its pharmacological properties.

The influential role of alcohol expectancies on men's post-drinking erectile response was identified with experiments using the balanced placebo design. It is a method for evaluating the independent and interactive influences of alcohol's physiological and psychological or expectancy effects. Two properties of drinking are manipulated systematically in the laboratory: expected alcohol content (expectancy) and actual alcohol content. Led to expect alcoholic or nonalcoholic drinks, half the research participants receive alcohol and half do not. Thus, the design has four groups: (1) Expect alcohol/receive alcohol, (2) expect alcohol/receive none, (3) expect no alcohol/receive none, and (4) expect no alcohol/receive alcohol. Use of drink "look/taste alikes" (e.g. vodka-tonic versus tonic or beer versus nonalcoholic beer) and limiting dosage to a low level ($<.05$ mg%) make the deception necessitated in groups 2 and 4 possible. Landmark experiments showed that men in the expect alcohol conditions exhibited greater erectile response compared to men in expect no alcohol conditions and actual alcohol content was inconsequential. The absence of an alcohol attenuation effect was likely due to the relatively low alcohol dosages necessitated by the balanced placebo design.

This expectancy enhancement effect seemed clear: Men who think they are drinking alcohol report and exhibit more erectile arousal than men who do not. The effect is best attributed to the self-fulfilling prophecy analysis outlined earlier. Also, consistent with this analysis, the expectancy enhancement effect on various sexual responses becomes more pronounced among men as a function of the strength of their *a priori* belief in alcohol's sexually enhancing or disinhibiting powers. Relatedly, survey studies using alcohol expectancy questionnaires have also shown that post-drinking sexual responding and sexual risk-taking are greater among men – and women – who more strongly endorsed faith in alcohol's aphrodisiac properties.

The expectancy enhancement effect was somewhat of a breakthrough in that the effect of "drink" on men's arousal could now be seen as consistent with the general

sexual disinhibition outcomes understood as commonplace in conventional wisdom. However, because people drink real drinks in the real world outside of the laboratory, there was an explanatory problem: The pharmacologically driven attenuation remained unreconciled with the anecdotal lore and empirical trends attesting to post-drinking sexual enhancement and disinhibition. This led to careful re-examination of the attenuation effect.

Critics noted that the attenuation effect was not universal. As additional studies accumulated, it emerged that attenuation effects were occurring in less than half of the relevant experiments. Also, in a study evaluating the highest dose reported to date (.15 mg%), alcohol attenuation of erectile arousal was not evident at all among men during sleep. Furthermore, methodological problems were identified, which may have artifactually fostered attenuation evidence in earlier studies.

In a recent set of experiments, these methodological problems were corrected and the attenuation effect was re-examined. Using men between the ages of 21 and 35 years recruited from the community, researchers evaluated the effects of two alcohol dosages (.08 and .10 mg%) in the context of instructions to control arousal by either maximizing or suppressing it while viewing an erotic film. At the lower dosage, alcohol did not attenuate genital arousal. However, alcohol interacted with arousal instructions to influence the latency to arousal onset. Specifically, sober men instructed to maximize arousal became aroused more quickly than did sober men instructed to suppress it. Intoxicated men did not show this distinction, suggesting that alcohol lessened their ability to control the onset of their arousal. At the higher dosage, alcohol did have an attenuating effect on men's peak genital arousal, but no effect on other erectile measures. In other related findings, three dosages (.06, 08, .10 mg%) were evaluated during exposure to an eroticized story depicting an attractive sexual situation and no alcohol attenuation effects were detected.

The most striking outcome of these recent findings is that, contrary to earlier attenuation findings, relatively high dosages of alcohol had limited impact on men's erectile responding. At the highest dosage, an alcohol attenuation effect did emerge, but it was confined to one of five measures. Specifically, alcohol reduced peak circumference change, but it did not affect average circumference change, latency to change onset, latency to peak change, or self-reported arousal. This circumscribed effect, when considered in the context of previous null effects at high dosages with wakeful and sleeping men, posed a challenge for the long held view that acute heavy drinking diminishes erectile performance. These data suggest that alcohol's attenuating effects necessitate higher blood alcohol content levels to manifest and are more topographically specific than previously thought – affecting mainly peak response. Also, alcohol's attenuating effects on erectile functioning were appearing to be more context dependent than previously recognized. A contextual factor that emerged as influential for response latency indices was the instructional demand to maximize versus suppress arousal. Generally, it was appearing that there was little basis here to view alcohol – in this range of acute dosages – as necessarily interfering with sexual response, for sexually functional men. It may be that the attenuation effect holds true only at the higher degrees of intoxication, much higher than .10%, which – while relatively rare in laboratory experimentation – may be common in men's real life encounters.

That said, an important consideration is limb of the blood alcohol curve, that is, whether the blood alcohol level is ascending or descending. In all of the above studies, except the sleep study, investigators sought to assess alcohol's effects on the ascending limb or peak portions of the blood alcohol curve rather than on the descending limb. Less is known about descending limb effects. Generally, ascending alcohol effects are characterized by heightened cognitive and motor impairment, euphoric mood, and stimulation; while descending effects are characterized by reduced cognitive and motor impairment, dysphoric mood, and fatigue. In the one study dedicated to examining alcohol's genital arousal effects on the descending limb among wakeful (rather than sleeping) men, an attenuating effect was evident at a moderate dosage, but again only in a narrowly confined way. Men in the alcohol condition (.08%) exhibited less genital arousal than no alcohol controls, but only when instructed to maximize arousal and only on one of three erectile measures. This finding indicated again that attenuation effects may be more limited than previously thought and also that the threshold for attenuation effects may be somewhat lower on the descending than ascending limb of the alcohol curve. The latter point is consistent with the idea that the descending limb could be especially prone to potentiate alcohol attenuation effects, because of the diffuse systemic inertia (dysphoria and fatigue) associated with descending blood alcohol. A potential implication of this finding is that in circumstances where men are deliberately exerting conscious effort to become maximally aroused, alcohol likely undermines peak erectile response for quite some time after drinking has ended and intoxication subsides. An interesting twist to these findings was that intoxicated men instructed to suppress their arousal exhibited more genital arousal than intoxicated men instructed to maximize their arousal. By implication, in circumstances where men are deliberately trying to resist becoming aroused, being on the descending alcohol limb may hamper their effectiveness at resistance. Reduced

effectiveness at suppressing arousal in conjunction with alcohol's tendency to heighten risk-taking could result in an increased likelihood of pursuing sexual activity in circumstances – such as nonconsensual or disease risk situations – where sexual pursuit should be curtailed.

Acute Episodic Erectile Failure

Consistent with the research finding that high levels of acute alcohol intoxication can attenuate erectile functioning, there are clinical and anecdotal reports of men experiencing erectile failure during episodes of heavy drinking. Phrases such as "whisky dick," "beer dick," and "brewer's droop" have emerged colloquially to characterize and label this phenomenon. While the causal mechanism would clearly seem to be physiological, there are significant clinical and psychological implications. First, frequent episodes of alcohol-induced erectile failure may signal the presence or development of alcohol abuse or alcohol dependence disorders and could serve as a prompt for seeking addiction evaluation and treatment. Second, even a single episode can contribute to the development of an erectile disorder. Specifically, if the episode is not accurately attributed to alcohol; but if instead it is misattributed to psychological concerns about virility, sexual potency, or partner issues then a recurrent pattern of erectile failure can develop and persist. Although initially caused by physiological processes, this pattern can become psychogenic in nature. For instance, subsequent alcohol-free sexual encounters can become characterized by partner conflict and performance anxiety about repeated erectile failure, both of which can drive emotional and cognitive distractibility that, in turn, disrupt normal erectile responding. Under these circumstances, alcohol moderation or abstinence and psychological and/or pharmacological treatment for erectile disorder may be advised.

Erectile Dysfunction Medications

A final note about alcohol and erectile functioning concerns medications used to treat erectile disorder. There is growing evidence that medications such as Viagra (Sildenafil) and Cialis (Tadalafil) are often used in conjunction with alcohol, both legally and illicitly. Among men diagnosed as suffering from erectile disorder, alcohol dependence or both, these medications appear to improve sexual functioning. Among sexually functioning nonalcoholic men, these medications are often used recreationally during drinking episodes, a pattern that has been especially observed in research findings with gay men. Little is known empirically about the interactive effects on sexual functioning of consuming alcohol with these drugs and research is warranted.

ALCOHOL AND ORGASM

Women

Approximately 25% of women report problems with orgasm to a clinical level. Overall, there is little research on alcohol and orgasm in women. To our knowledge, there is virtually no research examining the relationship between alcohol and problems with orgasm to a clinical level. The majority of studies, although few, examining the relationship between alcohol and orgasm has found that alcohol is related to more inconsistent orgasms or less orgasms overall in women. Heavy alcohol use, alcohol dependence, and alcohol use before sex are related to both higher rates of lack of orgasm and more inconsistent orgasms in women. In addition, alcohol abstinence increases the ability to achieve an orgasm. In contrast, one study using daily diaries did not support the finding that alcohol is related to orgasm in women.

To our knowledge, only one study has examined the effect of acute alcohol ingestion on orgasms in women. This study was conducted in the 1980s and had less than 20 participants, all of which were college students. It was found that latency to orgasm was longer and subjective intensity of orgasm was decreased for women with higher blood alcohol concentrations. However, it was also found that women with moderate and high levels of intoxication reported more pleasurable orgasms.

Men

Although 20–30% of men report orgasm problems indicative of premature ejaculation disorders, there has been practically no research on alcohol and male orgasm. In one study reported more than 30 years ago, sexually functional college men exhibited increased orgasm latency after drinking. Latency increased linearly with alcohol level (.03, .06, .09 mg%). Unlike with women, men's subjective orgasm ratings decreased with alcohol. Men reported less sexual arousal, orgasmic pleasure, and orgasm intensity with increasing alcohol dosage. However, the authors noted that this effect "may be viewed as a beneficial consequence" because "sexual intercourse would be prolonged for both partners."

SEXUAL PAIN IN WOMEN

There is very little research on the relationship between alcohol and sexual pain in women. The few studies that have examined this relationship are survey studies and the findings were consistent: alcohol use is

associated with more sexual pain. Also, one study found that clinical levels of sexual pain decreased with sobriety in a study of alcohol-dependent women. Because this research is only correlational, the direction of any causal connection is unclear: It is possible that women with sexual pain drink to ease the pain. However, there is not enough research to understand the relationship between alcohol use and sexual pain.

SUGGESTIONS FOR FUTURE RESEARCH

Important knowledge gaps remain concerning alcohol and sexual functioning. First, while sexual desire is a key construct in sexual functioning theory and clinical diagnostic scholarship, very little is known empirically about alcohol and sexual desire. Sexual desire is obviously an important motivational construct. A call for expanding research dedicated to elucidating how alcohol consumption affects this motivation and – reciprocally, how sexual desire can affect drinking – is clearly warranted. For instance, there are suggestions that people may sometimes drink alcohol as a way of heightening sexual desire or as a way of coping with anticipated sexual encounters. The latter is a pattern apparent among victims of sexual abuse and assault. However, very little is known about these phenomena. Second, regarding arousal, more work is needed specifying the physiological mechanisms accounting for alcohol's attenuation effects on arousal. Third, also regarding arousal, isolated experimental findings from the 1980s revealed that sexual stimulation and arousal can fuel increased drinking and suggested that further experimentation would be valuable. Fourth, little is known about alcohol and women's sexual pain disorders. Fifth, the preponderance of research on alcohol and sexual arousal has focused on men. Although research evidence related to alcohol and women's arousal has expanded exponentially in the last five years, many unaddressed questions remain to be formulated and pursued. Finally, the vast majority of research, especially experiments, on alcohol and sexual functioning have not considered the importance of variation due to culture, ethnicity, and sexual orientation. Research is needed to clarify whether the patterns observed among majority samples have full or qualified generalizability to minority populations, where stigmatization issues could play a role.

CONCLUSIONS

Alcohol's ubiquitous and durable association with sexual functioning will undoubtedly persist, as will sustained intellectual effort in understanding the nature of this relationship and its benefits and costs. Relevant science and clinical scholarship has provided a mix of impressionistic data, robust empirical trends, and recent clarifications to established conclusions. Impressionistically speaking, there is consensus that acute alcohol consumption in moderation enhances sexual desire, fosters greater arousal, and slows orgasmic response thereby prolonging intercourse. Empirically, several patterns are evident. First, heavy chronic alcohol consumption impairs sexual functioning and contributes to sexual dysfunctions, especially in men. Second, although there has been relatively little empirical research – especially experiments, acute alcohol consumption appears to be associated positively with sexual desire and related indices. Third, heavy acute consumption attenuates men's and women's genital responding. Also, in men, there are anecdotal and clinical indications that heavy acute consumption can cause episodic erectile failure, which can contribute to psychogenic erectile disorder. Fourth, contrary to earlier impressions and scientific conclusions, chronic and acute alcohol consumption in low to moderate dosages does not appear to impair sexual arousal. Regarding chronic consumption, recent analyses of population-based studies suggest that chronic non-heavy alcohol consumption may offer protection against erectile dysfunction. Regarding acute consumption, recent experiments have clarified that alcohol-induced attenuation of genital responding does not become reliably evident until the blood alcohol level exceeds .10 mg% on the ascending limb for both men and women or exceeds .08 mg% on the descending limb for men. Fifth, it is clear that alcohol's effects are not entirely physiologically driven and that alcohol's psychological properties – especially alcohol expectancies – play an important role in sexual functioning. Sixth, it seems clear that alcohol's role in sexual functioning directly and indirectly impacts other dimensions of sexual health, such as disease-related risk-taking. Regarding future research, as noted earlier, important knowledge gaps remain concerning alcohol and sexual functioning, especially concerning sexual desire and women's sexual arousal. As a final note, there has been insufficient attention to the benefits of alcohol's association with sexual functioning. With the understandable focus on problems and costs associated with alcohol-involved sexuality, the potential contributions of alcohol-involved sexuality to sexual health and overall well-being have gone largely unnoted and unexamined. As depicted in commonplace cultural images such as wedding night champagne, romantic dinner wine, and bikini beach tropical drinks, alcohol-involved sexuality can be a delight! This perspective, given its relation to human health and prosperity, also merits intellectual and research attention.

SEE ALSO

Alcohol and Sexual Violence, Alcohol, Sexual Risk Taking, and Sexually Transmitted Infections, Models of Relationships between Substance Use and Mental Disorders, Substance Use and Mood Disorders

Further Reading

Cheng, J., Ng, E., Chen, R., Ko, J., 2007. Alcohol consumption and erectile dysfunction: meta-analysis of population-based studies. International Journal of Impotence Research 19, 343–352.

Chew, K., Bremner, A., Stuckey, B., Earle, C., Jamrozik, K., 2009. Alcohol consumption and male erectile dysfunction: an unfounded reputation for risk? Journal of Sexual Medicine 6, 1386–1394.

Crowe, L.C., George, W.H., 1989. Alcohol and human sexuality: review and integration. Psychological Bulletin 105, 374–386.

George, W.H., Davis, K.C., Norris, J., Heiman, R.J., Schacht, R., Stoner, S.A., Hendershot, C.S., Kajumulo, K., 2011. Women's sexual arousal: effects of high alcohol dosages and self-control instructions. Hormones and Behavior.

George, W.H., Davis, K.C., Norris, J., Heiman, R.J., Schacht, R., Stoner, S.A., Kajumulo, K.F., 2006. Alcohol and erectile response: the effects of high dosage in the context of demands to maximize sexual arousal. Experimental & Clinical Psychopharmacology 14, 461–470. PMID: 17115874.

George, W.H., Stoner, S.A., 2000. Understanding alcohol and sexual behavior. Annual Review of Sex Research 11, 125–127.

Sobczak, J.A., 2009. Alcohol use and sexual function in women: a literature review. Journal of Addictions Nursing 20, 71–85.

CHAPTER

36

Substance Induced Myopia

Nora E. Noel, Jennifer A. Heaton, Brian P. Brown

University of North Carolina Wilmington, Wilmington, NC, USA

OUTLINE

DEFINITION

Substance Induced Myopia (or Substance Myopia) refers to a model often used to explain the acute effects of a psychoactive substance on a person's subsequent behavior in a particular situation. The idea is that when taken in a significant dose, the substance interferes with attentional processes and narrows a person's ability to perceive and comprehend all the stimuli in the environment (hence the term myopia used as an analogy to eye disorders). According to the myopia model, the higher the dose of a substance beyond a critical threshold, the more it interferes with information processing, leading the person to rely on only the most salient cues in the environment to interpret the situation. Generally, the model is referred to as alcohol myopia because the vast majority of the relevant research examines the effects of alcohol on behavior. However, from a hypothetical standpoint, intoxication with any psychoactive substance that limits attention could have a similar effect.

DEVELOPMENT OF THE MODEL

The term alcohol myopia was first coined by Claude Steele and Robert Josephs in a 1990 article in the *American Psychologist*. They were searching for an explanation for alcohol's inconsistent effects on disinhibition as well as a related factor, anxiety reduction. A commonly assumed mechanism for alcohol's association with risk-taking or engaging in socially unacceptable behavior is that intoxication temporarily interferes with one's inhibitions that might normally have kept these behaviors in check. The disinhibition hypothesis posits that under normal circumstances one is inhibited by anxiety from breaking social conventions or exposing oneself to danger. Alcohol intoxication is believed to reduce that anxiety, thereby disinhibiting the person's behavior. Much anecdotal evidence exists that show when people drink, they can become uncharacteristically nasty, unpleasant, aggressive, extroverted, sexual, and so on. As Tara MacDonald and her colleagues put it, common sense flies "out the window" and intoxicated people may engage in many risky behaviors that while sober they would not consider. Such behaviors include driving under the influence of alcohol or drugs, having sex without protection, or starting a fight with a large stranger. However, when tested under controlled conditions, such as in a laboratory, alcohol intoxication does not appear to disinhibit behavior consistently. Often, people do report less anxiety

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00036-X

when they are intoxicated, but sometimes, unpredictably, they do not. In fact, people may become even more depressed or anxious when drinking and, in some investigations, intoxication can be associated with more prudent or altruistic behavior than sobriety.

Likewise, a large amount of research on the tension-reducing effects of acute doses of alcohol (see the encyclopedia entry on the tension-reduction hypothesis) previously had shown that drinking alcohol can reduce tension or anxiety, but sometimes it does not. Tension reduction in response to alcohol intoxication is an unreliable and often unpredictable consequence. Moreover, as noted above, sometimes drinking even leads to increased tension, anxiety or depression, much to the disappointment of the person drinking for the purpose of tension reduction.

To explain the inconsistent effects of intoxication Steele and Josephs proposed a cognitive mechanism (they called alcohol myopia) which they said predicted alcohol's effects better than a disinhibition model. They defined alcohol myopia as "a state of shortsightedness in which superficially understood, immediate aspects of experience have a disproportionate influence on behavior and emotion, a state in which we can see the tree, albeit more dimly, but miss the forest altogether."

By way of explanation, an overarching principle of social cognitive research is that people's responses to a social situation are largely influenced by how they interpret the vast array of stimuli or cues they perceive in the situation. Given the same social situation, a person may make vastly different interpretations depending on the stimuli to which he or she attends. The person's response is then guided by their interpretation.

For example, suppose a man at a party becomes interested in a woman and attempts to strike up a conversation with her. She might react by ignoring him completely or even being hostile. Such a communication would be easy to interpret and he could move on to talk to someone else. On the other hand, if she responds to his overtures with conversation, smiles and laughter at his jokes, the interaction becomes ambiguous. Is she just being polite? Is she interested, but cautious? Is she attracted to him, too?

To answer these questions, he must observe her actions and the external situation to make a judgment about how he will respond to her. Obviously, it would be impossible for a person to allocate attention to every stimulus in an environment. Instead, the man focuses on his internal and the external cues in the situation that he deems important (or salient) stimuli and he filters out irrelevant cues. His internal state (he is interested in her) and his history of interacting with women may influence part of this filtering. Note that attentional filtering is a relatively automatic (non-conscious) ongoing process that happens rapidly (in milliseconds). Quite a bit of cognitive-psychological research demonstrates that even under ordinary circumstances people can miss what should be very obvious stimuli and events if they do not fit with their expectations. For example, inattentional blindness refers to the failure to perceive an event that would be quite noticeable while engaged in an attention-demanding task. The published scientific research is compelling, but the phenomenon is also famously demonstrated by the short video posted on You Tube in which the viewer is asked to count the number of ball passes a basketball team makes. Most viewers watch the team and at the end of the video seem completely unaware that a large man in a bear costume walked right through the center of the action. After the bear is pointed out, he is very obvious on the second viewing.

Thus, people are already selectively screening stimuli, often guided by what they expect, or want, to see. If we return to the man at the party, he is busy attending to social cues that could signal that the attractive woman is mutually attracted to him, unaware that he is selectively screening information he will use in his decisional process. He will use his filtered information to decide how to behave with her.

According to the alcohol myopia model, if the man is or becomes intoxicated, he will be cognitively impaired, thus having an even more limited capacity to process information. The selection filter becomes tighter, narrower, and he will be capable of attending to only the most salient cues, as determined by his internal state or, perhaps, some extremely obvious cues in the situation. He is more likely to miss other important information. Thus, if sober, he might notice that she is smiling (very evident or salient to him) but also be aware that subtly she is making eye contact with other people and appears to be edging away from him. She is probably not interested in conversing with him. However, if alcohol intoxication narrows his attentional processing (i.e. he experiences alcohol myopia), he is likely to interpret her social cues incorrectly, focusing only on her salient smiling behavior, and assume that she welcomes his company. His subsequent behavior with her might be socially inappropriate and very unwelcome.

To further illustrate how the myopic state might function, several researchers suggest that intoxicated men present a higher risk of sexual aggression than sober men because they focus only on their own internal sexual arousal, attending most closely to any positive behaviors exhibited by the woman, while failing to perceive or process any discouraging or negative behavior on her part. The latter cues are

usually less salient because they do not fit with his expectations.

BEHAVIORS THOUGHT TO BE AFFECTED BY ALCOHOL MYOPIA

Researchers have conceptualized alcohol myopia as the explanation for a fairly long and varied list of alcohol-related behavior. Initially, Steele and Josephs used the concept to explain alcohol's inconsistent effect on anxiety, hypothesizing that in a situation where there are no distracting or more salient stimuli, intoxication may cause people to focus on the obvious, salient anxiety-producing stimuli and become more anxious, rather than less anxious after drinking. Their experiments supported their hypothesis by showing that when intoxicated people focused on an impending anxiety-provoking event, they became more anxious, but, in contrast, when intoxicated participants were given a distracting task, more attention-demanding than the anxiety-provoking event, they became less anxious. Similarly, results of some observational and survey studies suggest that people may drink to escape or relieve depression (e.g. drown their sorrows), only to find themselves more depressed (e.g. crying in my beer) as they become more intoxicated. Thus, proponents of the alcohol myopia model have suggested that merely drinking, without engaging in an attention-grabbing distraction, leads to a focus on the person's internal state of depression, which then intensifies.

Other researchers have hypothesized that alcohol myopia plays a key role in alcohol-related aggression. Laboratory studies support the idea that alcohol consumption may increase aggression because the intoxicated individual focuses on reasons for fighting (e.g. "This person insulted me." "This person stepped ahead of me in line." "This person bumped me intentionally."). Reasons for fighting are often quite salient in a drinking situation such as a loud bar or party. According to the model, the alcohol-myopic person loses awareness of less obvious circumstances, both external (e.g. "That person has friends here.") and internal (e.g. "I don't want to fight with someone. I'd rather just have a quiet evening.") that might normally limit their aggression. He or she then reacts with aggression.

In a similar vein, many researchers suggest that alcohol-related sexual aggression increases when cues for the instigation of sexual behavior are more salient than cues inhibiting sex. According to the model, the intoxicated sexually aroused man might believe that he has the right to demand or force sex from his partner, perhaps because he thinks that the she instigated sex. Her sexual cues are salient to him and her protests are not. Moreover, several studies of risky sex, such as having sex with an unknown partner or without a condom, support the idea that when both parties are intoxicated they are likely to focus exclusively on their sexual arousal (the salient stimulus) and ignore less salient concerns such as becoming involved with an inappropriate partner, becoming pregnant, or acquiring a sexually transmitted disease. When sober, they might express this as, "What was I thinking?" An alcohol myopia model can provide a plausible *post hoc* explanation.

One set of alcohol myopia studies that is particularly interesting suggests a compelling explanation for driving under the influence while intoxicated with alcohol or other impairing substances. Sober college students who have been exposed to drunk-driving prevention campaigns often strongly endorse the belief that driving while impaired is irresponsible and dangerous. Even in anonymous surveys, many state that they intend to avoid drinking and driving. However, both laboratory and field studies show that when they actually are intoxicated, their responses change. Once their Blood Alcohol Level reaches a critical threshold (usually about 0.06%), they are much more likely to say that they intend to drive, particularly if their interviewer emphasizes that their intended destination (usually home) is not far away or follows a very familiar route. In such circumstances, the alcohol myopia model suggests that the desire to drive the car to an easy destination becomes paramount (salient) and the possible dangers (distant in time and likelihood) are discounted or ignored. An important further implication of this model is that it predicts (up to the point of incapacity) that the more intoxicated the person, the more likely he or she is to decide to drive while intoxicated. In fact, it predicts that the most intoxicated people are the ones most determined to drive themselves home.

Gambling, smoking, overeating, and other such risky or unhealthy behaviors while intoxicated have been studied as examples of alcohol myopia. Most of this work has been field studies or surveys and questionnaires, so the findings lack the experimental rigor that laboratory studies could have. Nonetheless, myopia seems such a universally applied and appealing explanation that the number and variety of intoxicated behaviors attributed to alcohol myopia continues to grow. Recently, one group of researchers even suggested that they often found association between suicidal behavior and intoxication might result from alcohol myopia intensifying suicidal or self-destructive thoughts.

DOES ALCOHOL MYOPIA EXPLAIN HOW INTOXICATION CAN LEAD TO MORE PRUDENT BEHAVIOR?

Demonstrating that intoxication is associated with poor decisions and discounting of negative consequences

is not, in itself, differential support for an alcohol myopia versus a disinhibition explanation. Evidence providing more compelling support for alcohol myopia comes from the studies mentioned above that show a counterintuitive increase of inhibitory effects associated with acute intoxication. A disinhibition model predicts increased risky behavior, regardless of environmental cues, whereas an alcohol myopia model differentially predicts decreased risky behavior if the most salient cues in the situation promote caution rather than risk.

As an example, one research group reported a series of experiments with a strong inhibitory cue (a hand stamp reading "AIDS Kills"). Volunteers had their hands stamped before they went out to socialize for the evening. Contrary to what most people would expect, it was the intoxicated participants whose hands were stamped before they went out who reported more safe sex practices than sober participants with the same hand stamp. In contrast, when participants had no hand stamp (or had a slightly less inhibitory one that read "Safe Sex") the results were reversed: sober participants were more likely than intoxicated participants to report using protective strategies. A disinhibition hypothesis would have predicted only that intoxication would lead to riskier behavior regardless of the hand stamp, so these results are more consistent with an alcohol myopia model. Hypothetically, if a person had a frightening, intense message (AIDS Kills) stamped on the back of a hand, intoxication caused such an intense focus on the message that they were more aware of it than of other external cues.

Other researchers have reported controlled studies in which intoxication was associated with more positive behavior than sobriety, again depending on which cues were most salient in the environment. Given the right stimuli, for example, intoxicated participants were more likely to be generous and helpful to others than sober participants. These results have all been interpreted as evidence that the intoxicated participants were focusing more on the emphasized very salient needs of others rather than less salient positive aspects of saving their money.

GOING BEYOND ALCOHOL: SUBSTANCE MYOPIA?

To date, although alcohol myopia appears quite embedded in the scientific literature, very few studies or papers address other psychoactive substances as a source of myopia. One reported laboratory study of nicotine effects was consistent with a nicotine myopia explanation, but otherwise, no controlled studies of any substance other than alcohol were found.

Expanding myopia research to encompass other substances could be a promising method of explaining differential use as well as differential acute effects of substance use. After all, alcohol myopia explanations of behavior are linked theoretically to attention allocation models in cognitive psychology (see the discussion of inattentional blindness above). Attention allocation models are well supported in the laboratory and have been applied to a much wider range of behavior than alcohol or Substance Induced Myopia. Substance myopia research might even be characterized as a subset of attention allocation research, thus arguing even more strongly for laboratory demonstrations of parallel effects and connections with more comprehensive models of perception and decision-making. So why are so few laboratory studies of substance myopia effects published?

One explanation for the lack of published substance myopia studies may be that they simply have not been tried. The cost of doing laboratory drug administration research certainly limits feasibility. Rigorous study of the acute effects of intoxication requires administration of the substance in a controlled setting in controlled doses. The expensive involved (e.g. ensuring participants' safety and comfort until the substance has dissipated) may limit researchers' ability, interest, and efforts to investigate these possibilities.

Of course, a second possible explanation is that studies with other substances have been done, but negative or non-significant results limit their likelihood of publication. This point will enter into the discussion below.

LIMITATIONS OF MYOPIA-BASED EXPLANATION

On the surface, alcohol myopia appears to have garnered much empirical support as an explanation of intoxicated behavior. However, this show of support may be misleading. Many of the studies cited as evidence for the myopia model are *post hoc* explanations of how associated factors may relate to one another. Findings that intoxicated people engage in risky and negative behaviors do not provide differential support for the alcohol myopia model versus the more parsimonious disinhibition explanation that it was intended to replace. The burden of support for the myopia model really falls on two types of studies: (1) those showing that counterintuitive behavior can occur during intoxication if the situational cues are manipulated and (2) those showing that distraction plays a key role in the behavior of acutely intoxicated people (e.g. the original studies by Steele and Josephs, described above). As noted above, such studies do not appear in great

abundance in the literature, so the support for the myopia model as a simple explanation is not as strong as many authors tend to suggest.

Further, a few published experiments, despite careful design and execution, have reported finding no evidence for myopia. At least one of these laboratory-based experiments examining sexual aggression suggests that a more complex model of disinhibition incorporating prior individual differences in participants' attitudes may better predict such behavior than does the alcohol myopia model (relying on situational cues). Critics of these experiments have suggested that perhaps the inhibitory cues used in the studies were not salient, but this is somewhat of a disingenuous and possibly circular argument. Alcohol myopia is actually a two-step process. First, certain cues must be salient or obvious enough to capture the participant's attention. Second, intoxication should result in a stronger focus on those cues versus others in the environment. As discussed above in reference to attention allocation, whether a cue is salient or not depends upon the eye, or rather the brain, of the beholder (i.e. the participant). A researcher cannot declare that certain cues were salient on the basis of how the participant behaved when under the influence of alcohol. Cue saliency is an important basic element of the model and should be assessed with the specific participants in every such experimental protocol. In fact, very few studies of the alcohol myopia model have reported a separate assessment of cue saliency. Interestingly, the experiment reported above that failed to find an alcohol myopia effect on sexual aggression did, in fact, report research demonstrating salience of the inhibitory cues with the target participant sample.

On a more general level, questions about natural cue saliency have been raised in support of the myopia model on the occasions whenever non-supportive results have been reported. One contention is that instigatory cues (e.g. cues to initiate sex and aggression) are naturally more salient than inhibitory cues and will always command more attention when they are present (no matter how intense the inhibitory cues may be). Again, because saliency is not often assessed separately in any of these experiments, further study is needed before such an argument can be accepted.

To date, the use of the myopia model to predict and to even modify antisocial types of behavior lacks support. A recent review used the principles of an alcohol myopia model to suggest methods of decreasing alcohol-related aggression incidents, but no empirical data were presented to support these possibilities. If the model is going to be useful beyond providing explanations of behavior, research on modifying behavior is sorely needed.

Finally, another limitation of myopia research is the ongoing definition and redefinition of intoxicated myopia. Currently, there are at least two variations of the model that have developed from studies of intoxicated aggression versus studies of the use of alcohol to relieve anxiety and depression. The two similar lines of research (both using the same terms) can lead to occasional confusion, especially when findings seem to fit with one model variation but not the other. Because all of the variations tend to rely on the same principles, we have presented this introduction to myopia as a relatively unitary model.

CONCLUSIONS

Substance Induced Myopia presents a cognitively based model of intoxicated behavior that has wide appeal, as evidenced by the number of times that it is cited *post hoc* to explain seemingly uncharacteristic actions of people when they are intoxicated. Alcohol myopia, in particular, has been used to understand intoxicated people's decisions to drive after drinking, engage in risky sex, behave aggressively toward others, do dangerous stunts, behave suicidally, become sexually assaultive, and so on. The major premise of the myopia model is that the more intoxicated they become, the more people will focus on the most salient cues or stimuli in the environment to guide their behavior.

Despite its wide acceptance, empirical support for the myopia model is not as strong as it initially appears. The best support comes from studies showing counterintuitive behavior when cues are manipulated and other studies showing that distraction is associated with relief of negative emotions (e.g. depression and anxiety) when a person is intoxicated. More research is needed especially to determine if the fundamental premises are supported (e.g. which cues are more salient than others? Are instigatory cues always more salient than inhibitory cues?). Additionally, research with psychoactive drugs beyond alcohol is necessary to establish if a myopia model is a more universal substance myopia model. Finally, and most importantly, research is needed to show that substance myopia models can reliably and accurately predict intoxicated behavior. Only then can it be useful in modifying the risky, dangerous, and unpleasant behaviors that it is meant to explain.

SEE ALSO

Binge Drinking, Cognitive Factors in Addictive Processes, Impulsivity, Disinhibition, and Risk Taking in Addiction

Glossary

Attention allocation an active (but often non-conscious) process through which people direct their focus to specific stimuli in their immediate environment. No one can possibly focus on all the stimuli that impinge on the senses, so certain stimuli are selected for focus.

Disinhibition inhibition describes a learned process through which a person is thought to restrain certain behavior, because the behavior might be considered socially inappropriate, potentially dangerous, hurtful to others, and so on. Consequently, disinhibition describes behavior hypothesized to occur when those restraints are lifted or temporarily disabled. A prominent explanation for alcohol and other drugs' effects on behavior is that the substances disinhibit a person's behavior.

Drug administration laboratory studies studies in which precisely controlled doses of drugs (including alcohol) are administered to human or non-human subjects for purposes of measuring a variety of responses to those drugs. Such studies are necessary to establish cause-and-effect relationships between the substance at different doses and the participant's behavioral responses. They are conducted in a controlled setting (usually a laboratory, although the laboratory might closely resemble a bar) in order to preserve the internal validity of the experiment by eliminating potential confounding variables. The controlled setting also protects the safety and welfare of participants by closely monitoring them until they are no longer intoxicated.

Field or observational studies (of substance use) studies in which behavior associated with substance use is observed in a setting more naturalistic than the controlled laboratory. Factors associated with the substance, such as those preceding, occurring concurrently with or following consumption of the substance can be assessed. One advantage of such studies is that they can provide input to help formulate specific hypotheses about drug effects that can subsequently be tested in the laboratory. A disadvantage is that such studies can only test the strength of associations because independent variables cannot be manipulated by the researcher.

Inattentional blindness the inability to "see" an object or event even though it is obvious when pointed out. Humans have a limited capacity for attention (see "attention allocation" above), so certain stimuli may not be perceived, even though they are in plain sight. Studies of inattentional blindness and associated studies of change blindness demonstrate that the mind's selection of attentional stimuli is a very active process, but most people are not aware of the process (i.e. it is non-conscious).

Inhibitory cues stimuli in the environment that signal the need for (or advisability of) restraining one's behavior and not acting upon impulse.

Instigatory cues stimuli in the environment that signal the likely availability of certain highly desired or highly rewarding consequences, such as tasty food and sexual satisfaction.

Intoxication a psychophysiological state induced by the ingestion of a psychoactive substance. Once the substance has crossed the blood–brain barrier, it may disrupt normal brain functioning in a variety of ways. Some of these disruptions appear to be pleasurable, at least initially, and probably drive the desire to consume more of the substance, even though some of the effects are harmful or potentially harmful. Intoxication is a vague term and has been used to describe only slight disruption of functioning (usually associated with a low dose of the substance) as well as extreme disruption.

***Post hoc* explanation** a hypothesis or set of hypotheses suggested "after the fact" that attempts to account for a particular event that has already happened. *Post hoc* explanations or hypotheses may be an interesting or intriguing method of generating testable hypotheses, but they capitalize on chance and cannot demonstrate cause-and-effect relationships. Such a demonstration requires a second step: using the *post hoc* hypothesis to generate and test an *a priori* hypothesis.

Tension-reduction hypothesis a prominent hypothesis regarding reasons for alcohol and other drug use that is broken down into two premises: (1) people use alcohol in order to relieve or reduce tension and (2) alcohol does, in fact, relieve or reduce tension. Two different kinds of studies are needed to test the two parts of the hypothesis. First, if a person is in an unpleasant state of tension (vaguely defined), will he or she drink more than someone who is not in a state of tension? Second, if a person is in a state of tension, can a dose of alcohol (versus a placebo) reduce or relieve that state? Empirical support for the hypothesis has been mixed, depending on how tension is defined, how dose is defined, how relief or reduction is demonstrated, and so on.

Salient (stimuli or cues) the stimuli in the environment most likely to capture the person's attention. Sometimes researchers define salient stimuli as the most obvious stimuli. However, given demonstrations of inattentional blindness, for example, salience might be difficult to define easily.

Further Reading

Giancola, P., Josephs, R., Parrott, D., Duke, A., 2010. Alcohol myopia revisited: clarifying aggression and other acts of disinhibition through a distorted lens. Perspectives on Psychological Science 5, 265–278.

Kassel, J., Unrod, M., 2000. Smoking, anxiety, and attention: support for the role of nicotine in attentionally mediated anxiolysis. Journal of Abnormal Psychology 109, 161–166.

Leonard, K., 1989. The impact of explicit aggressive and implicit nonaggressive cues on aggression in intoxicated and sober males. Personality and Social Psychology Bulletin 15, 390–400.

MacDonald, T., Fong, G., Zanna, M., Martineau, A., 2000. Alcohol myopia and condom use: can alcohol intoxication be associated with more prudent behavior? Journal of Personality and Social Psychology 78, 605–619.

Noel, N.E., Maisto, S.A., Johnson, J.D., Jackson, L.A., 2009. The effects of alcohol and cue salience on young men's acceptance of sexual aggression. Addictive Behaviors 34, 386–394.

Steele, C., Josephs, R., 1990. Alcohol myopia: its prized and dangerous effects. American Psychologist 45, 921–933.

CHAPTER

37

Metacognition in Substance Misuse

Marcantonio Spada[*], *Dawn Proctor*[$], *Gabriele Caselli*[**], *Esben Strodl*[$]

[*]London South Bank University, London, UK [$]Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, QLD, Australia [**]Studi Cognitivi, Modena, Italy

OUTLINE

THE METACOGNITIVE MODEL OF PSYCHOPATHOLOGY

Defining Metacognition

Metacognition is defined as "thinking about thinking." Flavell expanded this definition further, referring to the knowledge and processes involved in the appraisal, monitoring or control of cognition. The theory and application of metacognition evolved from developmental psychology but has since been applied across other domains, including ageing, memory, neuropsychology, and psychopathology. The great majority of theorists would agree in drawing a distinction between three basic aspects of metacognition: metacognitive knowledge, metacognitive experiences, and metacognitive control strategies. Metacognitive knowledge refers to the information that individuals hold about their own cognition and about strategies which impact on it. This knowledge provides a plan or guide for processing, the rules of which may be more (explicit) or less (implicit) amenable to conscious awareness and verbal expression. Metacognitive experiences involve the application of this knowledge to generate online appraisals and interpretations of specific mental events and processes. Metacognitive control strategies involve the execution of responses to control the activities of one's cognitive system.

The Self-regulatory Executive Function (S-REF) Model of Psychopathology

In 1996 Adrian Wells and Gerald Matthews proposed the self-regulatory executive function (S-REF) model (presented in Fig. 37.1), a metacognitive architecture

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00037-1

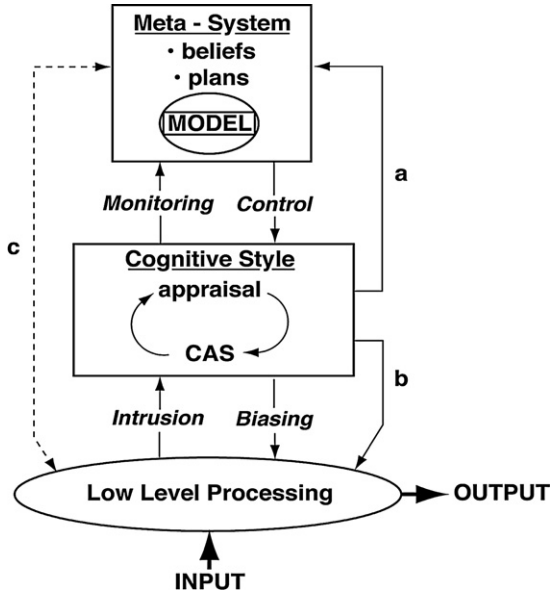


FIGURE 37.1 The S-REF model of psychological disorder. *Reproduced with permission from Wells, A. 2009. Metacognitive Therapy for Anxiety and Depression. Guilford Press, London, UK, p. 9 (Copyright 2009, The Guilford Press).*

for conceptualizing psychopathology based on three interacting levels of cognition.

In this model the first level consists of stimulus-driven processing networks which operate outside conscious awareness and can give rise to products which intrude into consciousness. Examples of these products may include affective (e.g. low mood), cognitive (e.g. negative thoughts), and physiological (e.g. palpitations) intrusions.

The second level of the model consists of the S-REF, an online, voluntary and conscious processing of actions and thoughts aimed at maintaining cognitive self-regulation in response to intrusions. Under adaptive conditions, S-REF activity is limited to short bursts. In these instances, an individual is able to select an appropriate strategy for either task-focused coping or modification of beliefs to successfully achieve adaptive cognitive self-regulation. However, when the individual is vulnerable to psychological disorder they may become locked in a cycle of processing known as the cognitive attentional syndrome (CAS). CAS configurations involve attentional bias toward threat, maladaptive coping behaviors, extended thinking (e.g. desire thinking, rumination, and worry) and thought control strategies. According to the S-REF model, these configurations are problematic because they cause negative thoughts and emotions to persist, and fail to modify dysfunctional self-beliefs, increasing the accessibility of negative information. The activation and persistence of the CAS in response to affective, cognitive, and physiological intrusions is dependent upon access to stored self-knowledge, which is contained in the third level of the model.

The self-knowledge at the third level of the S-REF model consists of metacognitive beliefs and associated plans for processing. Metacognitive beliefs are divided into two broad sets: (1) positive beliefs about coping strategies which impact on mental states, such as "Brooding will help me sort out things in my mind," or "Worrying will help me solve the problem"; and (2) negative beliefs concerning the significance, controllability and danger of particular types of thoughts, for example "It is bad to think thought X" or "I need to control thought X." The S-REF model broadly purports that positive metacognitive beliefs are linked with the activation of the CAS and negative metacognitive beliefs with its perseveration.

The S-REF model emphasizes the importance of the processes which generate, monitor, and maintain cognitions, rather than focusing upon the content of cognitions. Individuals are understood to select and implement coping plans based on metacognitive beliefs which may focus attention toward disorder congruent information or lead to the selection of unhelpful coping styles. This may establish a vicious cycle where the faulty blueprint is consistently applied to alleviate processes appraised as distressing but a successful resolution fails to be achieved. Rather than evaluating and updating the stored knowledge to adapt the blueprint accordingly, the individual continues to apply the faulty blueprint believing this will eventually facilitate cognitive self-regulation. Over time the combination of applying the same blueprint and failing to achieve a self-regulatory goal leads to the development of an internal dissonance characterized by negative appraisals toward the selected coping strategies and internal experiences more generally. Failure to achieve the goal may be linked to negative appraisals of the person's current state based on self-knowledge, unrealistic goals for cognitive self-regulation, external factors, or the application of inappropriate coping strategies.

APPLYING THE S-REF MODEL TO SUBSTANCE MISUSE

Recently, Marcantonio Spada and colleagues have applied the S-REF model to substance misuse. They proposed a triphasic formulation where metacognitive beliefs and the CAS operate across three phases: pre-substance use, substance use, and post-substance use. According to the formulation, the content of metacognitive beliefs and CAS configurations vary by phase,

person, and nature of the substance misuse presentation (e.g. severity, duration, and comorbidity).

The Pre-substance Use Phase

In the pre-substance use phase, individuals typically experience an intrusion (e.g. a craving, image, memory, or thought) that is associated with a strong positive or negative affective response, such as anxiety or excitement. The experience of this affective response is then appraised through the activation of metacognitive beliefs that in turn initiate the CAS with the aim of restoring arousal levels to a normative state. The activation of the CAS, however, leads to the escalation of negative affective responses and craving which in turn increase the likelihood of substance use initiation. Spada and colleagues argue that CAS configurations in this phase are predominantly characterized by extended thinking in the form of desire thinking and perseverative thinking (rumination and worry).

Desire thinking was described by David Kavanagh and colleagues as a voluntary cognitive process involving the elaboration of thoughts about a desired target: They particularly emphasized the motivational role of imagery. Recently, Gabriele Caselli and Marcantonio Spada made further distinctions between verbal (repetitive self-talk regarding the need to achieve the desired target and self-motivated statements) and imaginal (construction of mental images of the desired target or of its context of consumption) types of desire thinking. This thinking style has also been described as a preference or as a reaction to preference awareness. The target of desire thinking may be an activity, an object, or a state. Perseverative thinking is characterized by heightened self-focused attention involving persistent, recyclic and generic internal questioning regarding the causes, consequences and symptoms of one's negative affect (rumination) or attempts to engage in mental problem-solving on an issue with an uncertain outcome (worry).

A variety of studies undertaken by Spada and colleagues have demonstrated that negative metacognitive beliefs about the need to control thoughts and lack of cognitive confidence predict alcohol and nicotine misuse. Items relating to the "need for control thoughts" refer to beliefs that certain thoughts will lead to negative consequences unless they are not controlled. Items relating to "lack of cognitive confidence" represent metacognitive knowledge about the ineffectiveness of memory and judgment, and reflect diminished confidence in coping. The combination of these two forms of metacognitive knowledge may lead the individual to ruminate and worry about intrusions exacerbating negative affect and increasing the probability of substance use as a means of escaping the CAS and achieving, albeit temporarily, a degree of mental control. In support of this view research has demonstrated that rumination longitudinally predicts levels of alcohol use and category membership as a problem drinker. Several studies have also supported the association between high levels of worry and the tendency to use alcohol in order to reduce the worry process itself, especially among patients with alcohol use disorders.

A further cognitive strategy linked to the CAS which has been shown to increase negative affective responses and prolong extended thinking is thought suppression. Several studies have found that suppressing unwanted thoughts can lead to a paradoxical increase in these thoughts. Similarly, across substances such as tobacco and alcohol, attempt to control or suppress thoughts can lead to an increase in the frequency of unwanted intrusions and increased accessibility of substance-related information. The strategy therefore represents a counterproductive approach to reducing or managing substance-related urges or intrusions. Even where individuals report being able to effectively suppress unwanted thoughts this typically imposes an increased cognitive load that would detract from optimal cognitive performance when sustained over a prolonged period. Thought suppression has therefore been described as an unsuccessful attempt at coping with unwanted thoughts that typically increases psychological distress and therefore the likelihood of substance initiation.

In the pre-substance use phase the triphasic formulation emphasizes the importance of desire thinking in conjunction with the perseverative thinking and thought suppression as outlined above. Recent research by Caselli and Spada found that individuals who misuse substances hold both positive and negative metacognitive beliefs about desire thinking. Positive metacognitive beliefs concern the usefulness of desire thinking in controlling negative thoughts and emotion, in increasing positive sensations, in improving executive control over behavior, and in planning how to reach goals. These beliefs may be involved in the initiation of the desire thinking process as a form of coping with intrusions. Negative metacognitive beliefs concern the uncontrollability of desire thinking, and its negative impact on executive control over behavior, self-image and cognitive performance. These beliefs may play a role in propagating negative affect once a desire thinking episode has started, which in turn can contribute to the perseveration of desire thinking, leading to an increase in the sense of deprivation (as the substance target is elaborated upon but not achieved) and a greater likelihood of substance use initiation.

Depending on the individual and the substance of choice the activation of CAS configurations aimed at managing undesirable internal experiences may lead to escalations in cognitive–affective dysregulation and therefore increase the likelihood of substance use aimed

at restoring the normative state within the regulatory system.

The Substance Use Phase

In the substance use phase, positive metacognitive beliefs about substance use, which relate to the benefits of substance use as a cognitive–affective regulation strategy, are activated in response to an intrusion or following the activation of the CAS in the pre-substance use phase. These beliefs appear to play a central role in motivating individuals to engage in substance use. In addition to the activation of positive metacogn itive beliefs about substance use, a reduction in metacognitive monitoring appears to occur in this phase. Metacognitive monitoring refers to the monitoring of cognitive–affective change and proximity to goals (cognitive–affective regulation) as substance use proceeds. It relates to the individual's capacity for monitoring and reflecting upon cognitive processes to oversee the regulatory function of the S-REF system. Ineffective metacognitive monitoring reduces the flow of goal-progress information available for processing and simultaneously limits the opportunity to identify a stop signal that could guide action. For specific substances, such as alcohol, use itself will lead to a disruption in metacognitive monitoring. For example, an individual may hold positive metacognitive beliefs about the benefits of alcohol helping them to gain control over worrying thoughts. However, the alcohol impairs the individual's capacity to both reflect upon their internal state and monitor progress toward their goal of reducing their worrying thoughts. As this metacognitive monitoring is reduced, the person may or may not achieve their initial goal, but are unable to identify the stop signals that could guide them toward cessation which could result in prolonged use. During and following a substance use episode, this use is appraised as both uncontrollable and dangerous, contributing to an escalation of negative affect and triggering overuse of the substance.

In the S-REF model it is suggested that the central driver of this phase is the activation of positive metacognitive beliefs about substance use combined with reduced metacognitive monitoring, resulting in a pattern of behavior which is difficult to regulate as the individual loses sight of their objective (i.e. cognitive–affective regulation). While the activation of positive metacognitive beliefs about substance use contributes to triggering substance use, reduced metacognitive monitoring may prolong the substance use episode.

The Post-substance Use Phase

In the post-substance use phase the affective, cognitive, and physiological consequences of substance use are proposed to act as triggers for the activation of positive metacognitive beliefs about the benefits of cognitive reflection and post-event processing. The triphasic formulation suggests that such activation initiates perseverative thinking in the form of rumination and worry regarding substance use. As this process unfolds a growing number of intrusions may be generated which, in turn, activate metacognitive beliefs about the uncontrollability and danger of substance use thinking and related intrusions. This sequence may lead to escalations in craving and negative affect eventually bringing the individual to engage in further substance use as a means of self-regulation.

METACOGNITIVE THERAPY (MCT) FOR SUBSTANCE MISUSE

Adrian Wells and colleagues developed MCT as a transdiagnostic treatment approach for a range of psychological disorders. Empirical evidence supports the application of MCT for the treatment of generalized anxiety disorder, major depressive disorder, obsessive–compulsive disorder, and post-traumatic stress disorder. MCT consists primarily of a series of treatment components aimed at interrupting the CAS, including meta-level socialization procedures, shifting to a metacognitive mode of processing, modifying metacognitive beliefs, postponement of conceptual processes, attention modification (attention training technique (ATT) and situational attentional refocusing (SAR)), detached mindfulness and development of new plans for processing. The order of priority for delivering these treatment components varies in accordance to clinical presentations. What follows is an outline of MCT assessment and case formulation in substance misuse and examples of how the different treatment components of MCT can be used.

Assessment of Metacognition (AMC)

The primary purpose of the assessment phase is to gather information that will enable the development of an idiosyncratic case conceptualisation based on the triphasic formulation of the S-REF applied to substance misuse. Three forms of assessment that are relevant for substance misuse are outlined below:

The AMC Analysis

The AMC analysis provides a stepwise approach to assessment and case conceptualisation. The "A" represents an intrusion and this is followed by "M," the activation of a metacognitive plan. The plan consists of metacognitive beliefs and proceduralized blueprints

that guide coping and which may contribute to the CAS. In turn this leads to "C," the emotional consequences. In a brief example, the "A" is an intrusive thought about being socially inadequate. The metacognitive plan consists of (1) evaluating the thought as a significant intrusion that needs to be controlled; (2) worrying about whether the thought will persist; (3) activating beliefs about the benefits of using alcohol to control the thought; and (4) focusing on instrumental strategies for initiating substance use rather than monitoring internal states. The consequences include feeling anxious and ultimately using alcohol.

Metacognitive Profiling

The main goal of metacognitive profiling is to elicit the nature of the processing routines and metacognitive beliefs that are activated within the individual's metacognitive plan when faced with stressful or emotive situations. Metacognitive profiling may start by asking the client to describe a recent episode when they used a substance, paying particular attention to intrusions experienced before substance use was initiated. The assessment proceeds by focusing on metacognitive beliefs about the intrusion and about thought control processes. The nature of goals (typically cognitive–affective regulation) and cognitive processes that may play a role in achieving the goals (attention, memory, and judgment) are also explored.

Metacognitive-focused Psychometric Instruments

A variety of self-report instruments designed to assess metacognitive beliefs and aspects of the CAS typically aid the assessment process and facilitate the development of an idiosyncratic case formulation. These are summarized in Table 37.1.

Case Conceptualisation and Meta-level Socialization

The triphasic formulation of the S-REF model in substance misuse represents the basis for conceptualizing different CAS configurations and MCT pathways. Three key areas need to be specified across the triphasic spectrum: triggers, metacognitive beliefs, and CAS configurations. These key areas then need to be synthesized into a formulation of the client's substance use and fed back to them as a visual flowchart. The client will then have the opportunity to agree or disagree with the various components and modify them until a broad agreement is reached as to the formulation of the substance misuse.

The balance of focus on a given phase would depend on the severity and duration of the substance use presentation together with the client's level of awareness and treatment goals. For example, the pre-substance use phase may feature more prominently in occasional and irregular users or those who are in the early stages of

TABLE 37.1 Metacognitive-focused Psychometric Instruments

Instrument	Developers	Constructs assessed
Metacognitions Questionnaire (MCQ-30)	Wells and Cartwright-Hatton (2004)	Positive metacognitive beliefs about worry, negative metacognitive beliefs about the uncontrollability and danger of thoughts, cognitive confidence, negative metacognitive beliefs about the need to control thoughts and cognitive self-consciousness.
Positive alcohol metacognitions scale (PAMS) and negative alcohol metacognitions scale (NAMS)	Spada and Wells (2008)	Metacognitive beliefs pertaining to alcohol use.
Thought Control Questionnaire (TCQ)	Wells and Davies (1994)	Elements of the CAS, including strategies for controlling intrusions (distraction, punishment, reappraisal, social control, and worry).
Ruminative Responses Scale of the Response Style Questionnaire (RRS-RSQ)	Nolen-Hoeksema and Morrow (1991)	Rumination.
Penn State Worry Questionnaire (PSWQ)	Meyer, Miller, Metzger, and Borkovec (1990)	Worry.
Desire Thinking Questionnaire (DTQ)	Caselli and Spada (2011)	Imaginal and verbal aspects of desire thinking.
Alcohol Craving Experience Questionnaire (ACEQ)	Statham, Connor, Kavanagh et al. (2011)	Sensory aspects of craving (imagining taste, smell or sensations of drinking and intrusive cognitions associated with craving) when craving was maximal during the previous week (ACE-S: strength), and frequency of desire-related thoughts in the past week (ACE-F: frequency).

using to cope with negative effect. It may also play a central role for individuals attempting to remain abstinent but who have not addressed their pre-substance use mode of processing. In this sense it could represent a core theme for MCT focused on relapse prevention. The pre-substance use phase may have become highly "automated" in individuals with dependency and/or low levels of awareness. For these individuals the substance use phase would be the primary focus of formulation and treatment. Finally, the post-substance use phase may be of central importance when there is a chronic and persistent presentation with a history of relapses combined with a good awareness regarding the substance use problem.

Treatment Components

The overarching goal in MCT is to socialize the client to the role of the CAS in maintaining their current difficulties; specifically, in the case of substance misuse how: (1) desire thinking and perseverative thinking deepen distress and may lead to substance use; (2) metacognitive monitoring may lead to the perseveration of substance use; and (3) perseverative thinking about the substance increases negative effect and the possibility of further use. It is also important to highlight how the CAS prevents self-knowledge from being updated, particularly in the form of metacognitive beliefs.

Specific interventions need to be adapted to the principal target phase. All phases share the main treatment components (detached mindfulness, attention modification, modifying metacognitive beliefs and developing new plan for processing, defined below) but there are differences in each phase. Interventions for the pre-substance use phase need to focus on interrupting extended thinking, particularly desire thinking, and associated metacognitive beliefs. Interventions for the substance use phase need to primarily target attention modification (in particular the enhancement of metacognitive monitoring) and the challenging of positive and negative metacognitive beliefs about substance use. Developing new plans of processing to cope with negative affective states would be central to this phase as well as the pre-substance use phase. Finally, interventions for the post-substance use phase should promote a direct change in substance-related perseverative thinking (rumination and worry) and the modification of associated metacognitive beliefs.

Shifting to a Metacognitive Mode of Processing and Detached Mindfulness

This consists of developing new forms of awareness about cognitive events and processes. Clients are helped to see that the problem is not the occurrence of these events *per se* but their relationship to them, metacognitive appraisal of their impact upon the internal regulatory system, and strategies used to manage them. These internal events can range from craving and memories to thoughts, emotions, and physiological sensations. The detached mindfulness strategies are aimed at helping clients to move to an observing stance (or a metacognitive mode) with respect to their internal states. Detached mindfulness comprises two features: (1) mindfulness – an awareness of the internal experience; and (2) detachment – the suspension of any conceptual or coping activity in response to internal events and the separation of sense of self from the thought. Detached mindfulness is not a symptom management technique; it is intended to amplify the array of flexible responses to internal events. Detached mindfulness involves encouraging the client to observe their cravings, images, memories, and thoughts without trying to control or change them. These strategies are introduced using standard metaphors suggested by Adrian Wells and colleagues, such as taking the observer stance like watching clouds float across the sky or being a passenger on a train station watching trains go by, encouraging the client to relate to these cognitive events differently.

Attention Modification

This component involves the use of two strategies. In 1990 Adrian Wells developed the ATT, which aims to develop the individual's executive control over the allocation of their attentional resources. In terms of individuals with substance misuse, rather than automatically focusing their attention upon internal and external substance use cues, through ATT they may learn that they are able to choose to allocate their attention to non-substance related cues. The strategy involves asking the client to focus on a visual fixation point and to keep his or her gaze on this point. While doing this the client is directed to focus his or her attention on at least three competing sounds close by. The client's attention is moved between these sounds first slowly, and then rapidly and finally the client is encouraged to maintain attention on all of the sounds simultaneously. This pattern is repeated with at least three sounds outside the room and then at least three sounds far away. The clients are asked to practice this at least twice a day for 10–15 min.

The second strategy developed in 2000 by Wells involves SAR and aims to increase the flow of new and adaptive information into awareness so the individual is better able to update erroneous metacognitive beliefs (for example about uncontrollability). One application of SAR would entail encouraging the client to purposefully direct their attention onto substance cues and refrain from conceptual elaboration of substance-related pleasant memories and images. This could be combined

with fostering divided attention skills (e.g. extending the attention focus to include new and adaptive flow of information from the environment). Another application of SAR would involve enhancing metacognitive monitoring. The client could be asked to focus on the impact of substance use on internal states as a substance use episode unfolds. This scenario aims to provide a means of amplifying self-attention thus increasing the flow of goal-progress information into processing and increasing the possibility of substance use discontinuation.

Modifying Metacognitive Beliefs

In MCT applied to substance misuse, it is paramount to modify both negative and positive metacognitive beliefs. Negative metacognitive beliefs are usually tackled first as they tend to maintain the CAS. Examples of negative metacognitive beliefs may relate to various domains, including cognitive–affective states (e.g. "Anger is not good as it will lead me to lose control of my mind;" "My craving experiences are uncontrollable") or aspects of the CAS (maladaptive behavioral and thought control strategies –e.g. desire thinking, perseverative thinking, and substance use). Strategies for tackling negative metacognitive beliefs could include de-catastrophizing their significance through verbal reattribution (e.g. "What is the evidence for and against the idea that your cravings are uncontrollable?"), the facilitation of skills which promote a direct change in aspects of the CAS (for instance, interventions aimed at interrupting desire thinking, rumination, and worry) and practicing detached mindfulness. Positive metacognitive beliefs (e.g. "If I imagine the object I desire my mood will lift;" "If I ruminate I will find a solution;" "Using will help me gain control of my mind") are linked to the activation of unhelpful coping strategies. Modifying these beliefs starts with verbal reattribution such as an advantages–disadvantages analysis (e.g. "What are the advantages and disadvantages of thinking that drinking makes you gain control?"). This is followed by the exploration of better methods for achieving the advantages highlighted. Continuing on a similar theme, the client can be asked to explore the effectiveness of their chosen strategy in achieving their goal (e.g. cognitive–affective regulation).

In addition to these interventions, clients can develop a number of behavioral experiments to test their metacognitive beliefs. For example, clients may present with metacognitive belief where they believe that if they have an intrusion (e.g. craving) they need to act on it. A behavioral experiment could be devised by linking this metacognitive belief with a detached mindfulness activity. That is, clients can be exposed to the intrusion and then apply the detached mindfulness where they observe their internal state without acting on it. After this activity the clients are encouraged to reflect on their reaction to experiencing the intrusion.

Development of New Plans for Processing

Clients are encouraged to identify their use of coping strategies and then to not engage in them but to engage in healthy alternatives. This is done initially in session and then between sessions when clients experience internal states associated with use. For example, if clients are using distraction or thought suppression when they initially start experiencing unwanted intrusions, then rather than avoid these intrusions, they are encouraged to use detached mindfulness to just observe the experience without acting on it.

Substance misuse is seen as one of a number of behavioral coping strategies aimed at managing the individual's internal states. Clients are encouraged to identify their unhelpful behavioral coping strategies and then to not engage in them. For example, rather than avoid people or places which may trigger unwanted intrusions, clients will be encouraged to meet with these people and go to these places while using detached mindfulness or SAR.

SUMMARY

Metacognitive theory applied to psychopathology and substance misuse argues that the development and maintenance of problems may be linked to the monitoring, appraisal or control of internal cognitive processes. The S-REF model provides a conceptual framework to illustrate how stored knowledge and beliefs a person holds about their own thinking processes may influence the activation of plans for coping in response to cues or triggers from the environment. The emphasis of the model is upon the human drive to maintain cognitive regulation. In vulnerable individuals the configurations that are initiated to maintain this regulatory function can be unhelpful and maladaptive, leading to the development of a vicious circle consisting of negative patterns of responding which perpetuate cognitive–affective distress. Extending this model further, the recently developed triphasic formulation of substance misuse details a three-stage process that allows for the application of metacognitive theory to substance misuse.

During the pre-substance use phase the formulation suggests that cues trigger an affective response, leading to the activation of metacognitive beliefs regarding that internal state. The metacognitive beliefs guide the activation of the individual's blueprint for coping and this could include engagement in extended thinking (consisting of desire thinking and/or perseverative thinking), leading to the establishment of the vicious circle or CAS

and increasing the likelihood of cravings and substance use initiation. Due to this priming in phase 1 and the possible activation of positive metacognitive beliefs about the potential benefits of substance use for reestablishing internal regulatory control, the person may be more likely to shift to the second phase of the formulation. In addition, substance use simultaneously reduces the individual's capacity to track the regulatory state of the system by reducing metacognitive monitoring through the effects of the substance. In the post-substance use phase, the individual is proposed to select substance use perseverative thinking in response to the activation of positive metacognitive beliefs about the benefits of rumination and worry as post-event processing strategies. Perseverative processing of this nature is likely to lead to the activation of negative metacognitive beliefs regarding uncontrollability and danger as the person appraises further escalations in cognitive–affective distress. Strategies applied by the individual in this post-substance use phase could therefore increase the likelihood of future use to cope with ongoing distress.

MCT has shown to be effective in the treatment of various anxiety and mood disorders. The principles of metacognitive theory underpinning the triphasic formulation present a number of key components for therapeutic intervention that should be addressed according to the target phase. These include the interruption of extended thinking, challenging and restructuring of metacognitive beliefs, attention modification strategies, and the development of novel ways to manage and process negative internal states.

The triphasic formulation was developed based on empirical studies to date and offers a conceptual framework for the application of metacognitive theory to our understanding of substance misuse. Further research examining the interaction of metacognitive knowledge, plans for extended thinking and the relationship with internal and external triggers is required to consolidate this metacognitive framework. However, it offers a novel development that brings together our understanding of cognitive and affective distress in the development and maintenance of substance use disorders and provides a promising pathway for therapeutic intervention.

SEE ALSO

Cue Reactivity, Sensory Imagery in Craving, Deprivation, Craving, and Affect: Intersecting Constructs in Addiction, Attentional Biases in Craving

List of Abbreviations

AMC assessment of metacognition
ATT attention training technique
CAS cognitive attentional syndrome
SAR situational attentional refocusing
S-REF self-regulatory executive function

Glossary

Cognitive attentional syndrome transdiagnostic style of processing that is argued to be responsible for the maintenance of psychological disorder and consists of attentional bias toward threat, maladaptive coping behaviors, extended thinking (e.g. desire thinking, rumination, and worry) and thought control strategies.
Metacognition knowledge and cognitive processes involved in the appraisal, monitoring or control of thoughts.
Metacognitive control strategies aimed at controlling the activities of the automated system responsible for generating thoughts.
Metacognitive experiences application of metacognitive knowledge to generate online appraisals and interpretations of specific mental events and processes.
Metacognitive knowledge information individuals hold about their own cognition and about strategies which impact upon it.
Metacognitive monitoring process of maintaining self-awareness and reflecting upon internal states.

Further Reading

Caselli, G., Spada, M.M., 2010. Metacognitions in desire thinking: a preliminary investigation. Behavioural and Cognitive Psychotherapy 38, 629–637.
Caselli, G., Spada, M.M., 2011. The Desire Thinking Questionnaire: development and psychometric properties. Addictive Behaviors 36, 1061–1067.
Meyer, T.J., Miller, M.L., Metzger, R.L., Borkovec, T.D., 1990. Development and validation of the Penn State Worry Questionnaire. Behavior Research and Therapy 28, 487–495.
Nolen-Hoeksema, S., 1991. Responses to depression and their effects on the duration of depressive episodes. Journal of Abnormal Psychology 100, 569–582.
Spada, M.M., Wells, A., 2008. Metacognitive beliefs about alcohol use: development and validation of two self-report scales. Addictive Behaviors 33, 515–527.
Spada, M.M., Wells, A., 2009. A metacognitive model of problem drinking. Clinical Psychology and Psychotherapy 16, 383–393.
Spada, M.M., Wells, A., 2010. Metacognitions across the continuum of drinking behaviour. Personality and Individual Differences 49, 425–429.
Statham, D.J., Connor, J.P., Kavanagh, D.J., et al., 2011. Measuring alcohol craving: development of the Alcohol Craving Experience Questionnaire. Addiction 106, 1230–1238.
Wells, A., 2009. Metacognitive Therapy for Anxiety and Depression. Guilford Press, London, UK.
Wells, A., Cartwright-Hatton, S., 2004. A short form of the Metacognitions Questionnaire: properties of the MCQ-30. Behavior Research and Therapy 42, 385–396.
Wells, A., Davies, M.I., 1994. The Thought Control Questionnaire: a measure of individual differences in the control of unwanted thoughts. Behaviour Research and Therapy 32, 871–878.
Wells, A., Matthews, G., 1996. Modelling cognition in emotional disorder: the S-REF model. Behaviour Research and Therapy 34, 881–888.

Relevant Website

http://www.mct-institute.com/ – an international mental health company founded by Professors Adrian Wells and Hans Nordahl to offer metacognitive therapy resources and competency-based training workshops.

CHAPTER

38

Maturing Out

Andrew K. Littlefield, Rachel P. Winograd
Psychology Building, Columbia, MO, USA

OUTLINE

INTRODUCTION

Perhaps contrasting with popular notions, substance use disorders (SUDs) are largely disorders of young adulthood. Data from several large epidemiological studies indicate that the peak prevalence of SUDs, such as alcohol use disorders (AUDs), occurs during the early twenties and then drops precipitously with age. That is, many of the young adults who experience SUDs during young adulthood later remit (often without receiving treatment). This process of normative decline in problematic substance use is referred to as maturing out.

In this chapter, empirical data depicting the rise and fall in the prevalence of SUDs in young adulthood will be summarized. Specifically, theories describing potential mechanisms that contribute to the peak substance use observed during late adolescence/early adulthood will be described. In conjunction, theories that depict processes that contribute to maturing out will be covered. Finally, important considerations relevant to the phenomenon of maturing out will be highlighted.

DEFINITIONS OF SUDs

Before describing maturing out in more detail, it is important to note current conceptualizations of SUDs. The *Diagnostic and Statistical Manual*, version IV-TR, defines two major classes of SUDs: substance dependence and substance abuse. The diagnostic criteria for dependence is defined as "a maladaptive pattern of use, leading to clinically significant impairment or distress, as manifested by three (or more) of the

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00038-3

following symptoms occurring at any time in the same 12-month period: (1) tolerance; (2) withdrawal; (3) the substance is often taken in larger amounts or over a longer period than intended; (4) a persistent desire or unsuccessful efforts to cut down or control substance use; (5) a great deal of time is spent in activities necessary to obtain the substance, use the substance, or recover from its effects; (6) important social, occupational, or recreational activities are given up or reduced because of substance use; and (7) the substance use is continued despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by the substance" (p 197).

The DSM-IV defines substance abuse as "a maladaptive pattern of substance use leading to clinically significant impairment or distress as manifested by one (or more) of the following occurring within a 12-month period: (1) recurrent substance use resulting in a failure to fulfill major role obligations at work, school, or home; (2) recurrent substance use in situations in which it is physically hazardous; (3) recurrent substance-related legal problems; and (4) continued substance use despite having persistent or recurrent social or interpersonal problems caused or exacerbated by the effects of the substance" (p 198). These general criteria are applied to the diagnoses for specific substances (e.g. AUDs). According to the DSM-IV's diagnostic hierarchy, individuals who meet criteria for dependence for a given substance do not receive diagnoses for substance abuse, implying that dependence is the more severe of the two SUDs. However, for reasons that are beyond the scope of this chapter, the abuse/dependence distinction will most likely be eliminated in future versions of the DSM, resulting in the combination of the two classes of SUDs into one substance use disorder diagnosis.

DOCUMENTING MATURING OUT

Several studies have documented the peak prevalence of substance use and related negative consequences that occur during the early twenties before declining with age. For example, data from the Monitoring the Future study suggest that heavy drinking (i.e. drinking five or more drinks in a row in the past 2 weeks) peaks at ages 21–22 and then decreases linearly with age.

Numerous epidemiological studies have documented the rise and fall of SUDs during the third and fourth decades of life. Nationally representative data (for the United States) from the National Epidemiologic Study on Alcohol and Related Conditions (NESARC) clearly indicates that the peak prevalence of SUDs, such as AUDs, occurs during young adulthood and declines sharply with age. Notably, many of the individuals who recover from AUDs do not abstain from drinking but rather continue to consume alcohol at more moderate levels. Although some research suggests that heavy drinking among blacks and Hispanics may peak later compared to whites, and that the overall rates of AUDs differ among various racial-ethnic groups, the tendency for AUDs to decline with age is consistent among whites, blacks, Native Americans, Asians, and Hispanics. Similarly, NESARC data indicate that the past 12-month prevalence for drug use disorders (abuse and/or dependence for a broad range of substances including cannabis, cocaine, and heroin) is significantly higher in younger age cohorts (see Fig. 38.1). Thus, age-related declines in prevalence seem to be robust across substances and, at least in the United States, across racial-ethnic groups.

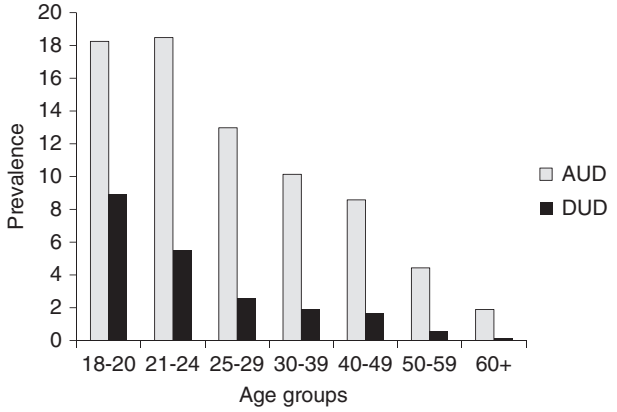


FIGURE 38.1 Prevalence of alcohol use disorders (AUDs) and drug use disorders (DUDs) from ages 18 to 60 in NESARC data.

THEORETICAL MODELS OF MATURING OUT

There are two essential questions regarding the phenomenon of maturing out: why does the peak prevalence of SUDs tend to occur in the early twenties, and why is there a dramatic normative decrease in both substance use and SUDs across the third and fourth decades of life? Potential explanations that have been provided by theorists and researchers are summarized below.

WHY DOES THE PEAK PREVALENCE OF SUDs OCCUR IN THE EARLY TWENTIES?

Several reasons have been proffered to explain the high rates of substance use and misuse that occur during the transition between late adolescence and adulthood.

Theorist Jeffery Arnett referred to the transition from late adolescence to adulthood as emerging adulthood, and describes it as a unique developmental period in which several factors contribute to problematic substance use.

According to Arnett, emerging adulthood lasts approximately from ages 18 to 25, though it can extend throughout the twenties and beyond for some individuals, depending on completion of developmental milestones such as marriage and parenthood. Arnett noted significant demographic changes over the past half-century in industrialized countries, such as prolonged educational pursuits and subsequent delayed entry into the workforce. This more recent focus on higher education could contribute to a later adoption of adult roles for some individuals. Indeed, compared to previous generations, both men and women are currently getting married and having children at later ages. Arnett suggests these demographic shifts result in the developmental period of emerging adulthood that is characterized by instability, self-focusing, feeling "in-between" adolescent and adult, increased optimism about the future, and identity exploration. As each one of these factors may contribute to delayed acquisition of adult roles, so too might they contribute to the increased use and misuse of substances.

For example, there are various forms of instability that are evident during emerging adulthood. As documented by Arnett's research, individuals frequently change residence, romantic partner, and educational and job status during emerging adulthood. As Arnett explains, this instability could result in anxiety and sadness, resulting in some emerging adults to use substances to "self-medicate" these negative emotions.

Emerging adulthood is also a time of exceptional self-focus. As opposed to adolescents (who are governed by parents and teachers) and adults (who have commitments to work and family), emerging adults have relatively few daily obligations to other people. As described by Arnett, self-focusing results in more volatile and ephemeral social networks and friendships, which in turn creates an overall lack of imposed social control compared to other age periods. This lack of social control coupled with relatively fewer responsibilities may also explain the heightened substance use during emerging adulthood.

Feeling "in-between" adolescence and adulthood is common during emerging adulthood. Arnett's research suggests that emerging adults consider reducing substance use and associated risk behaviors (e.g. drunk driving) as crucial criteria for adulthood, even more so than the adult transitions of graduating from college and getting married. However, since emerging adults feel "in-between" adolescence and adulthood, many may consider that this age period is a time where substance use is relatively acceptable compared to later age periods. In line with this perception of acceptability, research suggests that perceived norms of substance use is one of the strongest correlates of substance use behaviors among emerging adults. In fact, many emerging adults overestimate rates of peer substance use compared to their own use as well as the general acceptance of substance use and related behaviors. These beliefs may explain why many people engage in substance use during emerging adulthood but not later in life.

Arnett also describes emerging adulthood as the "age of possibilities," that is, a time period when people believe that they can make dramatic changes in their life and optimism for the future is exceptionally high. This optimistic bias, according to Arnett, may result in emerging adults to discount the myriad negative consequences associated with substance use, resulting in increased substance use during this age period. Supporting this conjecture, research suggests that many emerging adults base their decision to drink on perceived positive expectancies of substance use rather than potential risks. Further, Arnett hypothesizes that many emerging adults will have a high sense of well-being and use substances based on enthusiasm, whreas some emerging adults have a lower sense of well-being and use substance to cope with negative emotions. Supporting this notion, a large body of empirical literature suggests that alcohol use is increased among individuals who endorse relatively high levels of enhancement (e.g. I drink to have fun) and coping (e.g. I drink to relieve negative moods) motives to drink compared with individuals with lower levels of these motives. Further, alcohol problems are indirectly related to enhancement motives through increased alcohol, whereas coping motives directly predict alcohol-related problems. Thus, individuals who report higher motivations to drink tend to consume higher amounts of alcohol and subsequently experience more alcohol-related problems.

As Arnett describes, identity explorations may include emerging adults' desire to have a wide range of experiences before transitioning into adulthood, including experimenting with various substances. Further, distress related to forming a stable identity may result in some emerging adults to use substances as a means of coping with identity confusion (though there is limited existing empirical evidence to support this notion). Related to identity exploration, levels of sensation seeking (i.e. the pursuit of novel and intense experiences), a significant predictor of substance use, tend to be higher during emerging adulthood compared to later ages (see section "Why do People Tend to Mature Out of Substance Use?" for further discussion). Therefore, identity explorations that are common during emerging adulthood may contribute to the high prevalence of SUDs.

In line with Arnett's description of increased sensation seeking during emerging adulthood is a burgeoning collection of research that suggests overall mean levels of personality traits related to substance use are higher during emerging adulthood than perhaps during any other age period. For example, traits related to impulsivity (a broad construct that is related to making rash decisions and sensation seeking) and neuroticism (the tendency to experience negative emotional states) correspond to the increased use and misuse of a range of substances. For many individuals, these traits may be most pronounced during emerging adulthood and directly and indirectly relate to problematic substance involvement. For example, numerous studies have documented that individuals higher in traits related to impulsivity and neuroticism are more likely to have greater motivation to use alcohol. These motives in turn predict both alcohol use as well as alcohol-related problems. As will be described in more detail later in the chapter, changes in personality and motives across time may also influence the maturing out of substances.

In summary, the peak prevalence of substance use and SUDs occurs during emerging adulthood. This age period is marked by several factors that are conducive to substance use, such as increased freedom and exploration (compared to adolescence) and decreased responsibilities (compared to adulthood). Further, compared to older age groups, personality characteristics that are linked to substance use motives, substance use, and associated negative consequences seem to be relatively high during this timeframe. These features may help explain why substance use and misuse is the highest during emerging adulthood.

WHY DO PEOPLE TEND TO MATURE OUT OF SUBSTANCE USE?

As noted above, theorists and researchers have provided several explanations of the heightened use and misuse of substances that is observed during emerging adulthood. There is also a large body of work describing why there is a sharp decrease in substance use across the twenties and into the thirties. Primary findings from this literature are summarized below.

ADULT ROLE ATTAINMENT AND ROLE INCOMPATIBILITY THEORY

The most prominent explanation as to why many individuals mature out of problematic substance use is the attainment of adult roles. Numerous studies have documented that transitions into full-time employment, marriage, and parenthood are linked to decreased substance use and substance-related problems, especially alcohol use and misuse.

The link between adult role attainment and reduced problematic substance involvement may reflect selection or socialization processes. That is, the link between adult roles and maturing out might reflect the tendency of SUD remitters to later adopt adult roles (i.e. self-select). Contrastingly, adult roles may influence substance use through socialization processes, such as bringing about increased responsibilities, which result in substance use reduction for individuals who enter into these roles.

Work by Denise Kandel and Kazuo Yamaguchi suggests that both role selection and role socialization processes influence substance use. For example, findings suggest that marijuana use is associated with delayed entry into marriage and parenthood (i.e. a selection effect). Marriage and parenthood are also associated with increased rates of stopping marijuana use (i.e. a socialization effect). Additionally, using a longitudinal sample of college students, Kenneth Sher and colleagues concluded that the socialization effect of marriage on later AUDs seems to be more plausible than role selection effects of AUD on later marriage. Data taken from NESARC suggest that the relation between decreased alcohol dependence and getting married and having children reflect socialization processes, whereas the effects of school and work transitions were more consistent with selection effects.

Yamaguchi and Kandel proposed that the link between decreased substance involvement and adult roles could be attributed to role incompatibility, that is, adult roles associated with increased responsibility, such as marriage and parenthood, are "incompatible" with heavy substance use. For example, as Patrick O'Malley noted, marriage and parenthood result in numerous changes in social and recreational activities that are not conducive to maintaining a heavy drinking lifestyle. Specifically, compared to single individuals, those who are married attend fewer social functions that are compatible with substance use, such as parties and gatherings at bars. Increased responsibilities associated with child rearing also result in limited time for social activities involving substance use. Indeed, O'Malley noted that parenthood "is the key event" that propels men to reduce their drinking. Further, many women reduce their substance use during pregnancy because of concerns that substances such as alcohol can harm the developing fetus. Thus, many individuals seem to reduce their substance use in the presence of increased responsibilities and reduced free time that are brought about through the assumption of adult roles.

IS PERSONALITY CHANGE ASSOCIATED WITH MATURING OUT?

Although much of the research involving the maturing out of problematic substance involvement has focused on adult role obtainment, more recent work suggests that changes in personality may also be an important developmental factor. Despite personality being traditionally considered a stable, enduring construct, numerous studies have empirically documented mean level changes in personality across time and especially during emerging adulthood. These changes tend toward increased maturity, with people typically becoming less impulsive and more emotionally stable with age. This tendency has been referred to as the maturity principle. As with decreased substance use, increases in traits associated with maturity have been linked to adult role attainment and experiencing satisfaction in adult roles, such as marital and occupational satisfaction.

Recent work by Andrew Littlefield, Phillip Wood, and Kenneth Sher suggests that individual differences in personality change seem to influence the tendency to mature out of alcohol-related problems. That is, despite normative trends, there seems to be significant individual differences in both the tendency to mature out of problematic substance use and to experience increased psychological maturity. Thus, not all individuals show reductions in substance use and misuse or display beneficial changes in personality. Using longitudinal data from a cohort of college students, Littlefield and colleagues demonstrated that changes in neuroticism and impulsivity were linked to changes in alcohol problems from ages 18 to 35. Findings suggested that individuals who displayed sharper declines in neuroticism and impulsivity across this period also were more likely to undergo steeper decreases in alcohol problems. Subsequent analyses suggested these relations remained significantly linked even when adjusting for the influence of marriage and parenthood, demonstrating the unique importance of personality change in addition to the adoption of adult roles.

Using data from the same longitudinal sample, Littlefield and colleagues recently demonstrated that a group of individuals characterized by considerable decreases in impulsivity from ages 18 to 25 were more likely to show decreases in alcohol consumption and alcohol-related problems when compared to other groups with high and relatively stable levels of impulsivity. These findings suggest that individuals with relatively high levels of impulsivity who do not show significant decreases in that trait may be less likely to mature out of problematic alcohol involvement. These findings are consistent with recent longitudinal data suggesting that individuals who stopped smoking between the ages of 18 and 26 made corresponding larger increases in constraint (a construct considered to be the antithesis of impulsivity) and more pronounced decreases in negative emotionality compared to other individuals.

Individual differences in personality change also seem to influence changes in drinking motives, which in turn influence the tendency to experience alcohol-related problems. Individuals who exhibit more rapid decreases in impulsivity and neuroticism during emerging adulthood also seem to display larger decreases in coping motives to use alcohol. These reductions in coping motives predict more pronounced declines in alcohol-related problems. Thus, changes in personality and reasons for drinking seem to contribute to the maturing out of problematic alcohol involvement and most likely other substances as well.

Another potential, though somewhat speculative, explanation of the decline of substance use after emerging adulthood is the ongoing cognitive neurodevelopment that occurs during this timeframe. There is significant development in the prefrontal cortex during adolescence and emerging adulthood. These changes are thought to coincide with subsequent development of several processes, including affect regulation, self-control, cognitive control, and regulating behavior in general. Thus, declines in substance use after emerging adulthood, as well as the ostensibly adaptive personality changes detailed above, may be due in part to cognitive development.

In summary, the normative acquisition of roles associated with adulthood seems to strongly relate to the maturing out of substance use. This relationship likely reflects both role selection and role socialization processes. Just as the relatively lower amounts of responsibility and relatively higher amounts of freedom associated with emerging adulthood are conducive to substance use and misuse, the increasing responsibilities and time demands associated with adulthood seem to be largely incompatible with a heavy substance using lifestyle. Individuals also seem to undergo changes in personality during emerging adulthood, and changes in certain personality constructs seem to relate to changes in motives to use substances and decreased substance involvement. Normative cognitive neurodevelopment may also contribute to the maturing out of substance-related problems. In reality, changes in role status, personality, substance use motives, and cognition, as well as other factors, most likely affect each other and interact to produce dynamic, complicated influences on substance use behaviors (Table 38.1).

TABLE 38.1 Factors Potentially Contributing to Peak SUDs during Emerging Adulthood and Factors Potentially Contributing to Maturing Out

Factors potentially contributing to heightened substance use during emerging adulthood	Factors potentially contributing to declining substance use during early adulthood
• Increased freedom/relatively less responsibilities	• Adult role attainment/ increasing responsibilities
• Instability in roles, residence	• Role incompatibility
• Increased self-focus, decreased social demands	• Increased social demands
• Feeling "in-between" adolescence and adulthood	• Neurocognitive development
• Decreased focus on negative outcomes	• Decreased motives to use substances
• Identity exploration	• Identifying as an adult
• Social acceptance of substance use	• Decreased social acceptance of substance use
• Heightened levels of risky personality traits	• Decreases in risky personality traits

MATURING OUT: NORMATIVE BUT NOT UNIVERSAL

As described briefly above, there is considerable variability in the developmental course of SUDs. That is, although the normative trend is for rates of substance use and misuse to peak during the early twenties before declining with age, many individuals do not follow this typical pattern. For example, even among substance users, many individuals are not diagnosed with an SUD during emerging adulthood or later in life. Other individuals do not undergo normative decreases in problematic substance use but rather display chronic patterns of heavy use and SUDs across their lifespan. Still others experience SUDs and problematic substance use off and on at various intervals during their life. Thus, although the maturing out of substance use seems to be the norm (at least in the United States and other industrialized countries – see section "Cultural Differences" for further discussion), it by no means reflects the developmental course of substance use for all individuals.

ADULT ROLES DO NOT ALWAYS LEAD TO DECREASED SUBSTANCE INVOLVEMENT

Although transitions into adult roles seem to be a crucial element of maturing out, not all individuals who engage in adult roles reduce problematic substance involvement. Despite the harms associated with substance use during pregnancy, not all women quit (or even reduce) substance use habits when carrying a child; therefore, parenthood does result in reduced substance use in everyone. NESARC data suggest that entry into marriage is significantly associated with nonabstinent recovery (i.e. low-risk drinking) from alcohol dependence for many people. However, some individuals remained dependent after getting married and these individuals were less likely to subsequently recover from alcohol dependence. Additionally, though becoming a parent doubled the likelihood of becoming abstinent (i.e. eliminating alcohol use), this influence of parenthood was decreased among individuals who remained dependent after three or more years of starting full-time work. Therefore, not all individuals seem to mature out of SUDs after undergoing adult role transitions.

There are several reasons why adult role transitions may not always lead to reductions in substance involvement. Some occupations may be more conducive to substance use and misuse than others, and thus occupational obtainment may not always lead to decreased substance use. Occupational obtainment may also result in increased stress, which may result in increased substance use to cope with job-related stress. Positive assortative mating (the tendency for individuals to marry individuals with similar characteristics) may result in couples comprised of individuals who both engage in problematic substance use. Such relationships may not facilitate reduced substance use and, in some cases, may promote it.

For example, Kenneth Lenoard and his colleague Gregory Homish found that individuals with antisocial characteristics, a family history of alcoholism, higher levels of negative affect, and positive alcohol expectancies were more likely to engage in postmarital drinking. Drinking after marriage was influenced by partner drinking for both men and women. Decreased relationship quality (which may or may not be related to substance use) also seemed to predict later alcohol problems. The social network of the married couple may also moderate the influence of marriage on subsequent substance involvement. Leonard and Homish showed that the number of "drinking buddies" (i.e. social contacts that primarily involve activities associated with alcohol consumption) for both husbands and wives were longitudinally linked to both heavy drinking and alcohol problems. Taken together, these findings suggest that the characteristics of the individuals entering into marriage, as well as the nature of the marital relationship, moderate the protective influence of marriage on substance use.

There is also some evidence that the protective influence associated with adult role transitions may be

limited to certain age periods. Results from a meta-analysis of several studies suggested that never marrying was significantly related to increased alcohol consumption for men aged 18–39 but not older men. Findings also suggested that marriage/divorce was related to decreased/increased alcohol consumption in young but not older women. Thus, the specific age period in which adult transitions occur may alter the influence of adult roles on substance use.

DEVELOPMENTALLY LIMITED SUDs CAN STILL BE DESTRUCTIVE

Research documenting the phenomenon of maturing out suggests that many, if not most, individuals with SUDs during late adolescence and emerging adulthood later recover. Notably, this does not suggest developmentally limited SUDs are somehow benign. Extreme substance use, especially in the case of alcohol, is strongly linked to many of the injuries and deaths that occur during emerging adulthood. Many of the crimes that occur during this time period, such as vandalism, assault, and rape, are also alcohol related. SUDs that occur during emerging adulthood may also have detrimental influences on subsequent occupational and relational obtainment for the individual. Thus, though many individuals recover from SUDs, developmentally limited forms of pathological substance use still result in many negative consequences for societies, families, and individuals.

CULTURAL DIFFERENCES

Although most of the research described in this chapter is based largely on US samples, a meta-analysis of studies from ten different countries has suggested an international presence of maturing out. This pattern was consistent across most countries studied, indicating the cultural robustness of this normative trend. However, it should be noted that the countries included in this study (e.g. Germany, Poland, Switzerland, United Kingdom) are similar to the United States in their levels of Westernization/industrialization; thus, caution should be used when generalizing to other countries and cultures.

Just as legal drinking ages and drinking customs vary among regions, so too does the manner in which people develop and "grow out" of problematic use. For example, in countries with strict religious affiliations, often any amount of substance use is seen as destructive, whereas other regions that view drinking as an intrinsic part of the culture and heritage may have a more lenient stance on alcohol use. Thus, cultural and regional differences may affect the manner in which people mature out. Lastly, there is marked variability in the developmental course of SUDs in most, if not all, countries, with many people and subcultures not necessarily following the patterns and trends discussed throughout this chapter.

MATURING OUT IN EMERGING ADULTHOOD – HIGHER RATES OF RECOVERY OR DECREASED ONSET OF NEW CASES?

Throughout this chapter, the rapid decrease in prevalence of SUDs that occurs during the third decade of life has been discussed in terms of the high rates of recovery experienced by many emerging adults. However, intriguing new analyses from the NESARC data suggest that much of the rapid decline in the past 12-month prevalence of AUDs may be primarily attributable to a decrease in the rates of new onsets (i.e. new cases of AUDs) rather than increased offsets (i.e. recovery from AUDs) that occur during this time period. These findings suggest that SUDs, such as AUDs, may not be highly stable at any age; that is, recovery from SUDs is not limited to emerging adulthood. Rather, decreases in new cases of AUDs seem to occur more frequently in emerging adulthood than any other age group. Notably, these findings still suggest that AUD diagnoses are less persistent (i.e. a larger percentage of individuals with an AUD diagnosis later recover) during emerging adulthood than in other age periods. Overall, these results highlight the importance of understanding the many factors that contribute to acquisition, persistence, and recovery from SUDs during various age periods.

SUMMARY

For many individuals, emerging adulthood represents a time of peak use and misuse of various substances. Substance use during the time period relates to myriad negative consequences that could detrimentally influence life course. Fortunately, many, if not most, of these individuals later mature out of problematic substance use. Notably, these decreases in use typically do not occur in the presence of treatment. Importantly, many individuals continue to use certain substances, such as alcohol, though heavy consumption and negative consequences largely abate. These beneficial decreases are linked to adult role attainment as well as changes in certain personality traits.

Despite these normative trends, there is considerable variability in the developmental course of substance use and associated outcomes. Getting older is not a panacea

for problematic substance involvement. Unfortunately, some individuals continue to engage in risky patterns of consumption and experience negative consequences from substance use throughout their lives. Thus, understanding the contributing factors to risky and potentially chronic patterns of use at all life stages is of paramount importance.

SEE ALSO

Alcohol Use Disorders, Impulsivity, Disinhibition, and Risk Taking in Addiction, Interpersonal Factors and Addictive Disorders, Self-Medication, Adolescent Substance Use: Symptoms and Course, Symptoms and Course: Alcohol Use Disorder in Adulthood, Cultural Influences on Youth Alcohol and Drug Use, Costs and Consequences (Morbidity and Mortality) Associated with Adolescent and College Drinking and Related Problems

List of Abbreviations

AUDs	alcohol use disorders
NESARC	National Epidemiologic Study on Alcohol and Related Conditions
SUDs	substance use disorders

Glossary

Constraint a personality trait characterized by the tendency to reflect and deliberate before acting, and avoiding risky behaviors. Broadly, the opposite of impulsivity.

Emerging adulthood the transition from late adolescence to adulthood (roughly ages 18–25) that is marked by transitions, increased freedom, and identity explorations and in which peak rates of substance use and substance disorders are observed.

Epidemiological a scientific study that focuses on the causes, distribution, and spread of diseases or epidemics within a population.

Impulsivity a broad construct of personality that is characterized by the tendency of individuals to engage in behavior without adequate forethought about the plausible negative consequences of their actions. Broadly, the opposite of constraint.

Individual differences observed differences in individual characteristics and behaviors. For example, the fact that some people abstain from alcohol, whereas others drink heavily reflect individual differences in alcohol consumption.

Maturing out the process of a normative decline in peak problematic substance use that typically occurs across the mid- to late twenties.

Maturity principle the theory that people tend to undergo personality changes associated with increased maturity, such as becoming less impulsive and neurotic, thought to develop with the ability to handle the demands of adult roles.

Neuroticism a personality trait characterized by the enduring tendency to experience negative emotions and emotional states. Often translated to experiences of anger, guilt, anxiety, and depressed mood.

Substance use disorder use of alcohol or other drugs despite problems related to use of the substance. Substance abuse and substance dependence are the primary categories of substance use disorders.

Substance use motives an individual's reasons for using substances (drugs or alcohol). These include both positively reinforcing (e.g. drinking to enhance positive mood) and negatively reinforcing motives (e.g. drinking to cope with negative moods).

Role incompatibility theory a theory that describes the phenomenon when an individual's societal role (e.g. mother) is not compatible with that individual's behaviors (e.g. frequent drinking at bars).

Role selection the process in which individual differences influence the likelihood of assuming a specific role (e.g. emerging adult marijuana users are less likely to subsequently get married).

Role socialization the process in which the assumption of specific roles influence behaviors or other individual characteristics (e.g. marriage tends to result in decreased substance use).

Further Reading

Arnett, J.J., 2005. The developmental context of substance use in emerging adulthood. Journal of Drug Issues 35, 235–253.

Bachman, J.G., Wadsworth, K.N., O'Malley, P.M., et al., 1997. Smoking, Drinking, and Drug Use in Young Adulthood: The Impact of New Freedoms and New Responsibilities. Lawrence Erlbaum Associates, Mahwah, NJ.

Dawson, D.A., Grant, B.F., Stinson, F.S., et al., 2006. Maturing out of alcohol dependence: the impact of transitional life events. Journal of Studies on Alcohol 67, 195–203.

Jackson, K.M., Sher, K.J., Schulenberg, J., 2008. Conjoint developmental trajectories of young adult substance use. Alcoholism: Clinical and Experimental Research 32, 723–737.

Johnston, L.D., O'Malley, P.M., Bachman, J.G., Schulenberg, J.E., 2004. Monitoring the Future: National Survey Results on Drug Use. 1975–2003. In: College Students and Adults Ages 19–45, vol. II. National Institute on Drug Abuse, Bethesda, MD. NIH Pub No. 04-45508.

Johnstone, B.M., Leino, E.V., Ager, C.R., Ferrer, H., Fillmore, K.M., 1996. Determinants of life-course variation in the frequency of alcohol consumption: meta-analysis of studies from the Collaborative Alcohol-Related Longitudinal Project. Journal of Studies on Alcohol 57, 494–506.

Littlefield, A., Sher, K.J., Wood, P.K., 2010. Do changes in drinking motives mediate the relation between personality change and "maturing out" of problem drinking? Journal of Abnormal Psychology 119, 93–105.

O'Malley, P.M., 2004–2005. Maturing out of problematic alcohol use. Alcohol Research and Health 82, 202–204.

Raskin White, H., Jackson, K., 2004–2005. Social and psychological influences on emerging adult drinking behavior. Alcohol Research and Health 82, 182–190.

Sher, K.J., Gotham, H., 1999. Pathological alcohol involvement: a developmental disorder of young adulthood. Development and Psychopathology 11, 933–956.

Sobell, L.C., Ellingstand, T.P., Sobell, M.B., 2002. Natural recovery from alcohol and drug problems: methodological review of the research with suggestions for future readings. Addiction 95, 749–764.

Yamaguchi, K., Kandel, D.B., 1985. On the resolution of role incompatibility: life event history analysis of family roles and marijuana use. American Journal of Sociology 90, 1284–1325.

Relevant Websites

http://www.jeffreyarnett.com/ – Jeffrey Arnett's website – detailed discussion of "emerging adulthood."

http://www.niaaa.nih.gov – National Institute of Health – National Institute on Alcohol Abuse and Alcoholism, which includes a list of publications using NESARC data.

CHAPTER

39

The Effects of Substances on Driving

David J. Hanson

State University of New York, Potsdam, NY, USA

OUTLINE

SUBSTANCES AND DRIVING

Operating motor vehicles involves performing multiple tasks, the demands of which can change continuously. To operate a vehicle safely, a driver must remain alert, make decisions based on ever-changing information present in the environment, and execute maneuvers based on these decisions. This requires adequate visual functioning, information processing, decision-making, and psychomotor skills. The subtlety and complexity of the skills required to operate motor vehicles make them susceptible to impairment from a variety of both licit and illicit substances. However, less commonly recognized is the fact that some substances can enhance driving performance under certain circumstances.

The substances reviewed below are beverage alcohol (ethanol), caffeine, nicotine, cannabis (marijuana), amphetamine, cocaine, opioids (morphine and heroin), carisoprodol and meprobamate, phencyclidine (PCP), gamma-hydroxybutyrate (GHB) and its precursors of gamma-butyrolactone (GBL) and 1,4-BD, toluene, zolpidem (and zaleplon, zopiclone), dextromethorphan, diazepam, diphenhydramine, ketamine, and lysergic acid diethylamide (LSD).

SUBSTANCES

Ethanol

Ethanol, a central nervous system depressant, is a naturally occurring substance that results from the process of fermentation, by which yeast cells convert sugar into carbon dioxide and ethyl alcohol. Fermentation continues until all the sugar has been converted or until the fermenting mixture reaches about 14% by volume. At this point, the level of alcohol causes the yeast cells to die and the fermentation process to end.

Distillation is used to create alcoholic beverages of higher alcohol content or proof. In this process, heating causes the alcohol to vaporize before water; the alcohol vapors are then condensed into liquid ethanol.

Alcohol consumption is virtually always by oral ingestion. Ethanol vapors can be inhaled; however,

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00039-5

alcohol vaporizers are illegal in many jurisdictions and inhaling is a rare practice. It can also be absorbed transdermally, most often by means of enemas. Because both practices are probably unknown to most people, this discussion is limited to oral ingestion. After entering the mouth, a small amount of alcohol is absorbed into the tissues of the mouth and throat. Upon entering the stomach, it is absorbed into the bloodstream rather quickly because it does not need to be digested. Alcohol is a very small, water soluble molecule that quickly enters the bloodstream. About one-fifth of the alcohol consumed is absorbed in the stomach. The majority of alcohol is absorbed into the small intestine. However, a number of factors influence the rate of absorption, including the contents of the stomach. A large amount of food in the stomach dilutes the alcohol and slows its absorption there, and continuing to eat or nibble delays the release of stomach contents into the small intestine, where absorption is much more rapid.

Ethanol in the bloodstream (blood alcohol concentration (BAC), sometimes called blood alcohol level (BAL)) is measured by testing a sample of blood or is estimated by testing samples of exhaled breath. However, the latter method is subject to a number of measurement errors.

The results of either measurement or estimation of BAC are commonly expressed in terms of grams of absolute alcohol per deciliter of whole blood (g/dl). The g/dl notation is usually replaced by "percent" or "%," although it is not a true percentage because the numerical value represents a measure of weight in a measure of volume.

Metabolism occurs at a constant rate of about .015 of BAC per hour. If alcohol consumption exceeds the body's ability to metabolize it, the BAC will rise. After consumption ends, BAC will rise for a period of time before it begins to drop.

Men and women of the same height and weight who consume equal amounts of alcohol over the same length of time will have different BACs. Women have less of the gastric enzyme, alcohol dehydrogenase, which metabolizes about 20–30% of the alcohol in the stomach before it can be absorbed into the bloodstream. Women also have a higher proportion of body fat than men; alcohol is not soluble in fat and so it becomes more concentrated in the bloodstream of women than of men.

Driving-related skills are not all impaired at the same level of BAC. Wakefulness, for example, is generally impaired at lower BAC levels than simple reaction time and most components of visual functioning.

Although there is a fairly high correlation between a person's BAC and alcohol-induced impairment, there are large variations in the degree of impairment between individuals with the same BAC. One reason is that tolerance develops from repeated consumption of ethanol over a long period of time.

Generally unrecognized is the fact that impairment can also be influenced by a person's expectations. For example, impairment in performance of some tasks can occur if people falsely believe that they have been consuming alcohol. Relatedly, their impairment can be minimized if they falsely believe that they have not been consuming alcohol. The following section presents the results of numerous tests of the impairment or non-impairment of driving-related skills at different BAC levels.

Research has examined simple reaction time, which is measured by the length of time taken between a stimulus and the resulting intentional movement that a person makes. It is physiologically fixed in an individual and cannot be improved by practice. It can, however, be reduced by such factors as the ingestion of alcohol. Impairment of simple reaction time has been found at .043 BAC whereas some tests have found none at .080.

Choice reaction time is the time taken between a stimulus and an action that requires a choice. For example, a subject might be asked to give a response that corresponds to the stimulus, such as pressing a key corresponding to a letter if the letter appears on a screen or computer monitor. Tests of choice reaction time involve using multiple stimuli and multiple possible responses, thus requiring greater information processing and decision making than that required in simple reaction time experiments; both speed and accuracy are demanded. Unlike simple reaction time, choice reaction time can be improved by practice but can also be reduced by the ingestion of alcohol. It is more closely related to most actual driving tasks. Choice reaction time has been found to be impaired at a BAC of .020. However, it is more commonly found at BACs of .060 and higher.

Visual functioning includes eye movements, ocular motor control, and contrast sensitivity (ability to discern spatially distinct luminance differences). The brain's control of eye movements is highly vulnerable to the effects of alcohol. In driving, the eyes must focus briefly on important objects in the visual field and track them as they and the vehicle both move. Low BACs interfere with voluntary eye movements, which impairs the ability of eyes to rapidly track a moving target.

The frequency of eye movements and the duration of each fixation or "look" change significantly with increasing BAC levels. Research has examined the effects of alcohol on eye movement while subjects viewed a film of traffic events. The proportion of looks directed to the center of the driving scene increased under the influence of alcohol. As a result, subjects fail to see important peripheral events. Similar findings have been reported from on-the-road research. Impairment of visual functioning has been found at a BAC as low as .026.

Diverse psychomotor skills are required to operate a motor vehicle. Steering (tracking) is a complex psychomotor task in which the negative effects of alcohol on eye-to-hand reaction time are superimposed on the visual effects described above. Body balance and imbalance is also a commonly used indicator of psychomotor impairment by alcohol and is included in roadside sobriety tests. Impairment of psychomotor skills has been detected at .014 but is much more commonly found at BACs of .040 and higher. The driving-related skill of maintaining vigilance is impaired at moderate levels of BAC. It is regularly and consistently found at BACs of .030 and higher.

Alcohol impairs nearly every aspect of information processing by the brain. Alcohol-impaired drivers require more time to read a street sign or to respond to a traffic signal than unimpaired drivers; consequently, they tend to utilize fewer sources of information. Impairment of information processing is only occasionally found below .040 BAC.

Although wakefulness is not a skill, it is an essential condition for safe driving. Drowsy driving as a result of sleep deprivation is demonstrably a contributing factor to traffic crashes. Alcohol exacerbates sleepiness in its sedative effects as a depressant. This may be especially important because most alcohol-related crashes occur at night, which is a time when drivers are more likely to be sleepy. Wakefulness is impaired at BACs as low as .010 and always found at BACs of .035 and higher.

Caffeine

Caffeine is a naturally occurring central nervous system stimulant that temporarily reduces drowsiness and restores alertness. It is most often consumed in brewed beverages prepared from beans of the coffee plant (*Coffee arabica*) or the leaves of the tea bush (*Camellia sinensis*) as well as from various foods and beverages containing products derived from nuts (seeds) of the kola tree (*Sterculioideae cola*).

The drug is used medically to reduce fatigue, restore alertness, reduce drowsiness, and to treat problems sometimes found in premature infants. Caffeine is used recreationally for its taste, its stimulative effects, and to satisfy addiction to the drug. Over-the-counter caffeine tablets are often taken by students when studying for examinations, by people who work or drive for long periods of time, and by athletes to enhance their performance. Numerous studies have demonstrated often dramatic increases in endurance among athletes who consume caffeine. Caffeine has been shown to improve driving performance in actual road tests among drowsy drivers. There is also evidence that caffeine may increase alertness and improve reaction time after alcohol consumption but will not completely counteract alcohol impairment in a driver. Consumption of caffeine does not eliminate the need for sleep; it reduces drowsiness and increases alertness temporarily. This is important because maintaining alertness is important for safe driving. The National Highway Traffic Safety Administration has estimated that each year in the United States, about 100 000 motor vehicle crashes result from drowsy driving.

Nicotine

Nicotine is a naturally occurring central nervous system stimulant found in the leaves of the tobacco plant, *Nicotiana tabacum*. The primary therapeutic use of nicotine is in treating nicotine dependence in order to eliminate smoking and its attendant health risks, although it has many other therapeutic applications. However, the most common use of nicotine is recreational. Recreationally, nicotine is commonly inhaled with tobacco smoke, inhaled as snuff or in spray form, chewed in tobacco or chewing gum, inhaled as vapor (as in electronic cigarettes), placed between the lip and gum (as dipping tobacco or snuff), or absorbed transdermally through skin patches.

Nicotine has been shown to have a significant beneficial effect on athletic performance. It is similar in this regard to caffeine. Nicotine can both invigorate and relax users, depending on how much and how often they use the drug. When smokers want to achieve a stimulating effect, they tend to take short quick puffs, which produce a low level of blood nicotine concentration. When they want to relax, they tend to inhale deeply, which produce a high level of blood nicotine concentration that produces a mild sedative effect. Nicotine is unique in comparison to most other drugs in that it changes from a stimulant to a sedative as dosage increases.

Marijuana (cannabis)

Marijuana refers to the leaves and flowering tops of the hemp plant, *Cannabis sativa*. Cannabis contains substances called cannabinoids. These include tetrahydrocannabinol (THC), which is believed to cause most of the psychoactive effects of cannabis.

Marijuana is the most commonly used illicit drug throughout the world. It is usually smoked but can also be ingested orally. Data from traffic arrests and fatalities indicate that after alcohol, marijuana is the most frequently detected psychoactive substance among driving populations.

At typical recreational doses, the driving-related effects of cannabis include reduced inhibitions, disorientation, distorted perceptions of time and space, reduced

coordination, inability to concentrate, impaired memory, changes in thought formation and expression, drowsiness, and mood changes such as panic and paranoia.

The effects of marijuana vary to some degree with dose. However, it is difficult to establish a clear relationship between a person's THC blood concentration and its effects, including its impairing impact on driving-related skills. The commonly used indicator of psychomotor impairment by alcohol is included in roadside sobriety tests. The influence of any given blood THC concentration and specific effects is heavily influenced by the experience of the user, the individual vulnerability of the user to psychoactive reactions, the expectations of the user as to the effects of marijuana, and the setting of its use.

Although the onset of effects from ingesting marijuana orally take longer, the effects from smoking it are perceived by the user within minutes and peak after a period of between 10 and 30 min. The emotional high typically then lasts about 2 h. Most behavioral effects disappear within 3–5 h after use begins. Nevertheless, some researchers have found effects in some skills such as complex divided attention tasks for as long as 24 h. The more difficult and unpredictable the task, the more likely marijuana will impair performance. Impairment of psychomotor skills can continue after the emotional high has ended. The impairment of selective attention (the ability to disregard irrelevant information) among long-term users, even after abstinence, is adversely effected with increased length of use. Similarly, impairment in information processing speed is impaired with increased frequency of use. In laboratory studies, some drivers have demonstrated the ability to improve performance for brief periods of time by overcompensating for their self-perceived impairment. In real-life driving situations, the greater demands placed on drivers for long periods of time make impairment of driving performance a very serious traffic safety concern.

Combining alcohol consumption with the use of marijuana may produce dramatically greater impairment than does either substance on its own.

Amphetamine

Amphetamine is a powerful central nervous system stimulant, some of the effects of which can enhance the performance of driving-related skills. The drug can increase alertness, enhance the ability to concentrate, increase energy, and improve some cognitive functions including the ability to divide attention effectively.

However, amphetamine also creates tunnel vision, irritability, paranoia, aggressiveness, poor impulse control, psychomotor agitation, greater self-confidence, grandiosity, and strong feelings of power and superiority. These appear to be inconsistent with safe driving. The results of simulated driving after receiving very low doses of amphetamine (10–30 mg) have been inconsistent and may not be representative of driving performance at the doses ordinarily taken by recreational users (typically 100–1000 mg day^{-1}).

Because it reduces sleepiness, amphetamine can lead to insomnia, hours or days after which the user becomes exhausted and falls asleep (the "crash"). Among sleep-deprived persons on amphetamine, enhanced psychomotor skills have been found. However, the effects on cognitive performance appear to be mixed and inconsistent.

Amphetamine has long been, and is still, used by militaries around the world as a performance enhancer. British troops used 72 million amphetamine tablets during World War II and the Royal Air Force used so many that it was said that amphetamine won the Battle of Britain.

Amphetamine is also widely used as a physical and mental performance enhancer by civilians. Some students use it as a study and test-taking aid; it works by increasing energy levels, concentration, and motivation, thus allowing them to study for extended periods of time and to remain alert and concentrated in taking examinations.

Amphetamine is used by some high school, university, and professional athletes to enhance performance. It is also used by some truck and other commercial transportation drivers to increase energy, alertness, and concentration while driving for long periods of time. Like amphetamine, methamphetamine is a central nervous system stimulant. Possible driving-related effects of the drug include increased alertness, faster reaction time, poor coordination, aggression, poor impulse control, paranoia, delusions, and hallucinations.

Although it is sometimes used in the treatment of narcolepsy, attention deficit hyperactivity disorder, and obesity for a period of up to 6 weeks, it is not used for longer periods because of the serious risk of addiction and abuse. Tolerance may develop and users may quickly become addicted and use it with increasing frequency and in increasing doses.

Recreationally, methamphetamine is also used to enhance performance, create intense euphoria, relieve fatigue, increase alertness, self-medicate for depression, and control weight. Users often begin with intranasal or oral use, progress to intravenous use, and occasionally to smoking. Methamphetamine is sometimes used with alcohol or marijuana, especially during withdrawal. Peak blood methamphetamine concentrations occur shortly after intravenous injection, a few minutes after smoking, and several hours after oral ingestion. Effects are less intense after oral ingestion than after smoking or injecting.

Ecstasy is a "designer drug" that is synthesized to produce effects similar to those of amphetamine. Although it is a weaker central nervous system stimulant than amphetamine, it can cause muscular rigidity, tremor, decreases in attention, impaired visual functioning, and reduced ability to divide attention effectively in driving-related tasks.

It is difficult to predict the overall driving-related effects of ecstasy use. Driving simulation tests have found it to be associated with increased speed and more variation in speed. However, some tasks are not influenced whereas others are improved. Lacking in research is attention to the effects of different levels of dosing.

Cocaine

Cocaine is a strong central nervous system stimulant that is sometimes used as a topical local anesthetic for ear, nose, and throat surgery or other medical procedures. Recreationally, it is used to create euphoria, increase alertness, reduce fatigue, and create feelings of well-being and strength.

Cocaine also increases mental focus, mental clarity, and improved performance of some simple tasks. At higher doses it can cause confusion, disorientation, delusions, hallucinations, fear, paranoia, aggression, and antisocial behavior.

Use of cocaine can minimize the adverse effects on performance caused by sleep deprivation and also by the consumption of alcohol. Among people who are not sleep deprived, some research finds no effect on cognitive or psychomotor performance whereas others find enhancement of psychomotor performance, attention, and learning.

Specific effects are not closely associated with specific blood cocaine concentrations because of individual levels of tolerance and other factors. However, the onset of effects is slower and last longer than those of amphetamine.

Opioids (Morphine and Heroin)

Morphine is a naturally occurring substance extracted from the seedpod of the poppy plant, *Papavar somniferum*. Heroin is, in turn, produced from morphine. Both are central nervous system depressants.

Morphine is used medicinally for both acute and chronic pain management and is sometimes used to sedate patients. Heroin can be used as an analgesic and cough suppressant although these are no longer accepted medical uses in the United States and some other countries.

The effects of morphine or heroin depend on the dose, the route of administration (of which there are numerous possibilities), and tolerance. Both drugs can create driving-related effects including euphoria, inability to concentrate, drowsiness, slower reaction time, distractibility, mental clouding, tremors, lethargy, apathy, and depressed consciousness.

The effects of a dose of morphine typically begin within 15 min–1 h and last for 4–6 h. It may cause sedation and significant psychomotor impairment for up to 4 h following a single dose in normal individuals.

The driving performance of cancer patients who received long-term morphine for its analgesic benefits was not considered to be impaired enough by its sedative effects to increase the risk of crashes. Compared with a control group, the patients treated with morphine showed no impairment in vigilance, concentration, divided attention, or psychomotor skills. A small decrease in reaction time was found 3 h after administration of morphine. There are several field reports of driving impairment (weaving, poor vehicle control, poor coordination, slow response to stimuli, delayed reactions, difficultly in following instructions, and falling asleep at the wheel) in instances of drivers who have tested positive for morphine in law enforcement cases of driving while under the influence. However, the blood amphetamine levels are not reported nor is the existence of other drugs in the blood of the impaired drivers indicated.

For heroin, intense euphoria lasts from about 45 s to several minutes, strong effects last about 1–2 h, and then dissipate in 3–5 h. Slowed reaction times up to 4 h in former narcotic addicts have been found in laboratory research. As with amphetamine, high levels of tolerance can develop making effects less pronounced for the same dose in long-term users. Alcohol increases the effects of morphine, including sedation, drowsiness, and decreased motor skills.

Carisoprodol and Meprobamate

Carisoprodol and meprobamate are both central nervous system depressants that are legally available only by prescription. The former is prescribed to relax muscles and the latter is prescribed to reduce anxiety; both are dispensed in tablet form. The driving-related skills affected by either include loss of coordination, sluggish movements, tremor, reduced visual functioning, confusion, disorientation, slowed cognition, and lack of comprehension

The effects of carisoprodol begin within 30 min of ingestion and continue for as long as 4–6 h. The effects of meprobamate last much longer. Single doses of carisoprodol (700 mg) do not appear to effect cognitive or psychomotor performance. However, single doses of meprobamate can impair divided attention, slow reflexes, increase reaction time, and impair coordination.

Signs of cognitive and psychomotor impairment in persons found to be driving under the influence of either carisoprodol or meprobamate include poor perception, impaired reaction time, confusion, disorientation, inattentiveness, slurred speech, slow responses, sleepiness, lack of coordination, and difficulty standing, walking or exiting vehicles.

Similarly, among persons involved in driving under the influence cases who were found to be positive for carisoprodol and/or meprobamate and in which no other drugs were detected, impairment included slow reflexes, disorientation, sleepiness, poor balance, poor coordination, and slurred speech. Driving behaviors that were observed included extreme weaving, striking other vehicles and fixed objects, and hit-and-run accidents of which the driver appeared to be unaware. The greatest impairment was found among those whose combined use of carisoprodol and meprobamate led to high blood concentrations of those drugs.

Alcohol increases the sedation and mental confusion that can be produced by carisoprodol. Alcohol also increases the sleepiness, disorientation, incoherence, and confusion that meprobamate can cause. The use of other central nervous system depressants can also contribute to impairment.

PCP

PCP is a synthetic chemical that is illegally produced or illegally obtained from veterinary sources. It is currently used as a veterinary anesthetic or tranquilizer and used recreationally as a psychedelic and hallucinogen. Recreationally it is smoked, orally ingested, inhaled, injected, absorbed transdermally, and applied in the form of eye drops.

A single dose of PCP appears to create severe impairment of driving-related skills. The effects, which tend to be dose dependent, include blurred vision, impairment of eye–hand coordination, memory impairment, disorientation, disordered thinking, distorted sensory perceptions including distorted perceptions of space, memory impairment, poor concentration, hallucination, combativeness, feelings of power and invincibility, paranoia, and numbness in extremities. Effects generally occur within 1 h after dosage and slowed reaction time has been reported for up to 14 h.

GHB as well as GBL and the Drug 1,4-BD

GHB is a central nervous system depressant that is used medically as a sedative, an anesthetic, a treatment for narcolepsy, and to suppress symptoms of alcohol-dependence withdrawal and opiate withdrawal syndrome.

Recreationally, GHB is used for its effects in creating euphoria, reduced inhibitions, and sedation, and by bodybuilders as a substitute for anabolic steroids.

GHB was first synthesized in 1960 and sold as a food and dietary supplement. Thirty years later, the US Food and Drug Administration (FDA) banned the nonprescription sale of the drug. However, it is easy to produce illegally and ingredient kits with instructions are sold.

Later the FDA issued warnings about the dangers of GBL and 1,4-BD, both of which are used as industrial solvents and as ingredients in a variety of cleaners, paint removers, and degreasers. GBL and 1,4-BD are both rapidly converted into GHB in the body and are used as GHB substitutes. They are marketed as natural diet supplements, “antiaging” drugs, mood enhancers, energizers, as aids in achieving weight loss, and as self-medication for insomnia, anxiety, and depression.

Low doses of GHB do not impair attention, alertness, vigilance, psychomotor coordination, short-term memory, nor increase the effects of low BACs. However, the doses given in laboratory studies have been below those used recreationally.

Driving-related effects at higher doses include visual disturbances, loss of peripheral vision, confusion, reduced inhibitions, drowsiness, combativeness, short-term amnesia, shaking or seizures, and hallucinations. Signs of impaired performance in 24 cases of driving under the influence in which GHB was detected included drowsiness, confusion and disorientation, reduced peripheral vision, incoherent speech, short-term memory loss, lack of balance and unsteady gait, poor coordination, and poor performance of field sobriety tests. The impaired drivers had typically been stopped because of erratic driving, including such things as weaving, ignoring road signs, and near-crashes. The existence of drugs for which the drivers were not tested might have exacerbated the effects of the GHB.

Whereas the effects of amphetamine begin within 10–20 min, peak within 20–45 min, and generally last 2–5 h, there is a longer duration of effects following 1,4-BD ingestion because it metabolizes more slowly to GHB than does GBL.

Toluene

Toluene is a volatile solvent that acts as a central nervous system depressant. It occurs naturally in the tolu tree (*Myroxylon toluiferum*) and crude oil. It is produced during the refining process to make gasoline and other fuels from crude oil, in making coke from coal, and in manufacturing styrene. Toluene is used as a solvent in paints, lacquers, thinners, glues, and nail polish remover. It is also used in industrial processes such as printing and tanning leather.

Toluene has no medical use but is frequently used recreationally for its intoxicating effects. Such use is common among younger adolescents because the product is readily available, inexpensive, and legal, although some states have prohibited its sale to minors.

The driving-related effects of toluene include drowsiness, reduced ability to concentrate, slowed reaction time, distorted perception of time and space, impaired color vision, impaired vigilance, confusion, memory loss, delusions, and hallucinations.

No simple relationship between blood toluene concentrations and degree of impairment has been found.

Zolpidem (and Zaleplon, Zopiclone)

Zolpidem is a central nervous system depressant that is prescribed as a short-term treatment for insomnia. Zaleplon and zopiclone are also used for this purpose. The driving-related effects of each of these three drugs are drowsiness, double vision, slow reflexes, poor coordination, confusion, decreased cognition, difficulty concentrating, memory impairment, and hallucinations.

Effects typically last as long as 4–5 h. However, some effects can last for 8–16 h after heavy doses of any of the drugs. Given the same dosage, zaleplon has a more rapid onset and shorter duration of effects compared with zolpidem, whereas zopiclone has a longer duration of effects.

The recommended dose of zolpidem is 10 mg. Following doses of 10–20 mg, confusion, disorientation, double vision, cognitive impairment (memory, time estimation, memory), and psychomotor impairment can occur. Some of the impairments can last as long as 8 h. A 10–20 mg dose of zaleplon can cause significant impairment for 1–3 h with no effects on actual driving within 5–10 h. A 7.5 mg dose of zopiclone can cause severe effects on actual driving for as long as 10 h.

In five reported cases of driving impairment in which zolpidem was the only drug detected, behavior included erratic driving, slow reflexes, disorientation, confusion, loss of balance and coordination, loss of short-term memory, double vision, poor attention, an inability to stand or walk unassisted, and poor performance on field sobriety tests. In six other cases of driving under the influence of zolpidem, drivers were involved in automobile crashes or drove erratically, and symptoms included slow and slurred speech, unsteady gait, confusion, and disorientation.

Dextromethorphan

Available both by prescription and over-the-counter, dextromethorphan is used for the temporary relief of coughs caused by minor throat and bronchial irritation. It is used recreationally for its ability to create euphoria, elevated mood, dissociation of mind from body, dream-like experiences, and heightened perceptual awareness.

Although adverse effects with recommended doses are rare, slight drowsiness and dizziness can occur. Driving-related effects of higher doses for recreational purposes can include confusion, impaired judgment, psychomotor impairment, disorientation, reduced memory altered perception of time, and both visual and auditory hallucinations.

Combining dextromethorphan with promethazine may cause marked drowsiness or impair the mental or physical abilities necessary to operate a motor vehicle safely.

Diazepam

Diazepam is a central nervous system depressant that has numerous medical uses. Commonly prescribed doses are 5–40 mg daily.

Diazepam is used recreationally as a sedative or to enhance the effects of alcohol or opioids. For example, administration of diazepam 30 min after a dose of oral methadone produces a greater high. Diazepam is also used by cocaine users to increase the threshold at which seizures occur, by heroin users to enhance the effects of heroin, and by both cocaine and heroin users to reduce withdrawal symptoms between doses.

At low doses, diazepam is a moderate tranquilizer causing sleepiness, drowsiness, confusion, and some loss of memory. At high recreational doses, diazepam can produce slurred speech, blurred or double vision, slowed reflexes, disorientation, tremor, and hallucination. Alcohol consumption increases some of the effects of diazepam including drowsiness, sedation, decreased motor skills, and impaired memory.

A single dose of 5–20 mg of diazepam can cause significant impairment in driving-related skills with the greatest effects occurring at approximately 2 h after dosing and lasting for as long as over 4 h. Reported have been decreased attention, reduced divided attention, increased lane crossing, reduced auditory and visual reaction time, increased braking time, decreased eye–hand coordination, and impairment of memory, tracking, vigilance, information retrieval, psychomotor, and cognitive skills. Consumption of small amounts of alcohol additionally reduces impairment.

In drug users, diazepam can lead to greater impairments in psychomotor and cognitive performance, reduced concentration, and amnesia; some effects can last for a period of days. Epidemiological studies have reported a range of increased levels of risks to driving from taking diazepam. These increases have ranged

from twice to several times higher when compared with drug free drivers. The elderly may be at additionally increased risk of a vehicle crash.

Regular use of diazepam will produce tolerance to most of the sedative and adverse effects. Therefore, blood diazepam concentrations are not highly correlated with behavioral effects.

Diphenhydramine

Diphenhydramine is a central nervous system depressant that is used for the relief of allergy symptoms, for cough relief, as a sleep aid, and for the prevention of motion sickness.

The effects of diphenhydramine are increased by alcohol, monoamine oxidase inhibitors (MAOIs), diazepam, hypnotics, sedatives, tranquilizers, and other central nervous system depressants. Alcohol enhances such effects as drowsiness, sedation, and decreased motor skills. Diphenhydramine decreases alertness, reduces vigilance, decreases reaction time, induces drowsiness, reduces working memory, impairs concentration, reduces divided attention, impairs the estimation of time, impairs tracking, decreases learning ability, and reduces psychomotor performance. It impairs tracking, reaction time, and overall driving performance in actual on-the-road tests.

A single therapeutic dose of diphenhydramine can significantly impair psychomotor performance for 4 h and has been found to cause a greater negative impact on driving performance than the consumption of alcohol at a level exceeding the maximum legal BAC limit.

Ketamine

Ketamine is a hallucinogen and dissociative anesthetic that is legally available only by prescription, usually as a veterinary tranquilizer and anesthetic. It is also used occasionally as a short-acting general anesthetic for children and elderly patients.

The drug is used recreationally as a psychedelic and for its dissociative effects. Users have reported that the physical effects of ketamine are similar to those of PCP and that the visual effects are similar to those of LSD. Because it is difficult to synthesize, the ketamine that is used recreationally is usually stolen or illegally diverted from legitimate sources.

Ketamine causes impaired cognition, disorientation, blurred vision, increased distractibility, lack of coordination, increased reaction time, hallucinations, and distorted perceptions of surroundings, time, space, and sound. Effects can last as long as 2 h.

LSD

LSD is a hallucinogen and psychedelic drug that is manufactured from the lysergic acid that occurs naturally in the ergot fungus that grows on wheat and rye. The liquid is typically applied to blotter paper squares, stickers, sugar cubes, candy, or soda crackers. LSD is also available in dropper bottles and in the form of gelatin sheets. There is no medicinal use for LSD; it is used recreationally as a hallucinogen and for its ability to alter perception and mood.

LSD creates hallucinations, reduced reaction time, diminished visual acuity, altered mental state, thought difficulties, delusion, temporary psychosis, and distorted time and space perception. Effects begin within 20–30 min, peak at 2–4 h, and gradually diminish over 6–8 h. Some effects may last longer. Flashbacks may occur suddenly, often without warning, and may occur within a few days or more than a year after use. The effects of LSD are unpredictable and depend on the dose, the user's personality and mood, expectations, and surroundings.

Research suggests that the incidence of LSD in driving under the influence cases is extremely rare.

CONCLUSION

In a recent nation-wide roadside survey in the United States, 11.0% of daytime drivers and 14.4% of nighttime drivers tested positive for drugs, both medicinal and illegal. In the same survey, 0.9% of daytime drivers and 11.2% of nighttime drivers tested positive for alcohol.

It is important to note that testing positive for either alcohol or drugs does not necessarily indicate impairment. However, the evidence does suggest that the problem of drugged driving may be more serious than generally recognized.

Because of its legality, long tradition of use, social acceptability, and general availability, ethanol has long been widely recognized for its ability to impair driving performance. Law enforcement of impaired driving is almost exclusively focused on impairment from alcohol consumption. Law enforcement officers typically do not have the expertise or equipment necessary to test for the presence of other substances. Nor is there strong social or political pressure to do so. Consequently, some drug users report having been repeatedly stopped while driving under the influence of drugs but being released because they did not test positive for alcohol. By not consuming alcohol before driving, drugged drivers can operate their vehicles with virtual impunity. This problem will continue unless there is a significant change in

public perceptions. For over 30 years mothers against drunk driving (MADD) has been a strong leader in raising public awareness of the serious nature of drunken driving, an effective rallyer of moral indignation, a powerful force pushing for stricter laws, a persistent advocate for the strict enforcement of alcohol-impaired driving laws, and a vigilant monitor pressing for punishment of drunken drivers to the fullest extent of the law. However, MADD has simultaneously and consistently opposed any consideration of such traffic safety problems as drugged driving and driving while using a cell phone, which it sees as potentially competing for public attention and funding.

The problem of traffic crashes and the injury and deaths they cause is a very serious one. Dangerous driving has many causes. It includes not only alcohol consumption but also drug use, cell phone use, driving while drowsy, driving while eating, and many other factors, all of which need to be addressed.

A person who is killed in an impaired driving crash is just as dead whether the impairment was caused by alcohol, other substances, or driver inattention.

List of Abbreviations

BAC	blood alcohol concentration
FDA	US Food and Drug Administration
GBL	gamma-butyrolactone
GHB	gamma-hydroxybutyrate
LSD	lysergic acid diethylamide
MADD	mothers against drunk driving
PCP	phencyclidine
THC	tetrahydrocannabinol

Glossary

Blood alcohol concentration (BAC) Blood alcohol concentration, sometimes called blood alcohol level or BAL, is commonly expressed in terms of grams of absolute alcohol per deciliter of whole blood (g dl^{-1}) and measured by testing a sample of blood or is estimated by testing samples of exhaled breath.

Choice reaction time Choice reaction time is the time taken between a stimulus and an action that requires a choice.

Distillation Distillation uses heat to create alcoholic beverages of higher alcohol content or proof, causing the alcohol to vaporize before water; the alcohol vapors are then condensed into concentrated liquid ethanol.

Simple reaction time Simple reaction time is the length of time between a stimulus and the resulting intentional movement that a person makes.

Visual functioning Visual functioning includes eye movements, ocular motor control, and contrast sensitivity (ability to discern spatially distinct luminance differences).

Further Reading

Baselt, R.C., 2001. Drug Effects on Psychomotor Performance. Biomedical Publications, Foster City, CA.

Burns, M., Page, T., Leikin, J., 1998. Drug Information Handbook for the Criminal Justice Professional. Lexi-Comp Inc, Hudson, OH.

Couper, F.J., Logan, B.K., 2001. GHB and driving impairment. Journal of Forensic Sciences 46, 919–923.

de Gier, J.J., Hart, B.J., Nelemans, F.A., Bergman, H., 1981. Psychomotor performance and real driving performance of outpatients receiving diazepam. Psychopharmacology 73, 340–344.

European Monitoring Centre for Drugs and Drug Addiction, 2007. Drugs and Driving. Office for Official Publications of the European Communities, Luxembourg.

Galski, T., Williams, J.B., Ehle, H.T., 2000. Effects of opioids on driving ability. European Respiratory Journal 15, 590–595.

Ghoneim, M.M., Hinrichs, J.V., Mewaldt, S.P., Peterson, R.C., 1985. Ketamine: behavioral effects in subanesthetic doses. Journal of Clinical Psychopharmacology 5, 70–77.

Gjerde, H., Smith-Kielland, A., Normann, P.T., Morland, J., 1990. Driving under the influence of toluene. Forensic Science International 44, 77–83.

Hurst, P.M., 1987. Amphetamines and driving. Alcohol, Drugs and Driving 3, 13–16.

Joó, S., 1994. Methadone substitution and driver ability: research findings and conclusions from a discussion of experts. Journal of Traffic Medicine 22, 101–103.

Klonoff, H., 1974. Marijuana and driving in real-life situations. Science 186, 317–324.

Logan, B.K., 1996. Methamphetamine and driving impairment. Journal of Forensic Sciences 41, 457–464.

Logan, B.K., Couper, F.J., 2001. Zolpidem and driving impairment. Journal of Forensic Sciences 46, 105–110.

Moeller, M.R., Hartung, M., 1997. Ecstasy and related substances – serum levels in impaired drivers. Journal of Analytical Toxicology 21, 591.

Moskowitz, H., Fiorentino, D., 2000. A Review of the Literature on the Effects of Low Doses of Alcohol on Driving-Related Skills. National Highway Traffic Safety Administration, Washington, DC.

Physicians' Desk Reference. 2000. Medical Economics Company, Montvale, NJ.

Siegel, R., 1987. Cocaine use and driving behavior. Alcohol, Drugs and Driving 3, 1–7.

Stephens, B.G., Baselt, R.C., 1994. Driving under the influence of GHB? Journal of Analytical Toxicology 18, 357–358.

Vainio, A., Ollila, J., Matikainen, E., Rosenberg, P., Kalso, E., 1995. Driving ability in cancer patients receiving long-term morphine analgesia. Lancet 346, 667–670.

Relevant Websites

www.nhtsa.gov – National Highway Traffic Safety Administration (Drug Impaired Driving)

www.drugabuse.gov – National Institute on Drug Abuse (Drugged Driving).

www.whitehousedrugpolicy.gov – Office of National Drug Control Policy (Drugged Driving).

www.druglibrary.org – Schaffer Library of Drug Policy (References on Drugs and Driving).

CHAPTER

40

Defining and Assessing Drug Craving

Harold Rosenberg

Bowling Green State University, Bowling Green, OH, USA

OUTLINE

ADVANTAGES OF ASSESSING CRAVING

The experience of craving is commonly reported by substance abusing and dependent smokers, drinkers, and drug takers and is often a target of assessment and intervention in clinical settings. For example, a recent survey of American substance abuse treatment agency directors revealed that clients' experiences of craving are assessed regularly during both intake evaluations and ongoing therapy. Most respondents viewed craving as an important target of treatment, which their agencies addressed during both individual and group therapy sessions, often by providing education about craving, teaching coping skills, and encouraging clients to avoid or leave situations where craving occurs. Almost half of the respondents said their agencies would refer clients for anticraving medications, but only small proportions of respondents indicated that they administered published questionnaires designed to assess craving or offered cue exposure therapy in which clients practice coping with craving in the presence of drug paraphernalia or the sight and smell of their target drug.

Researchers are also interested in being able to assess the subjective experience of urges or craving. For example, researchers depend on being able to manipulate and assess craving to test predictions based on psychological theories of craving such as those summarized below. In addition, craving has been assessed as an outcome variable in the evaluation of psychotherapies such as cue exposure therapy and pharmacological therapies such as drug-replacement medications and purported anticraving medications. Both clinicians and researchers have studied whether craving predicts a drug taker's likelihood of lapse or relapse some days, weeks, or months after assessment.

Recognizing the advantages of being able to assess craving, a variety of methods – including self-reports of subjective craving, self-reports of other behaviors considered indicators of craving, and observation of

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00040-1

behavioral and physiological indicators of craving – have been employed to assess craving for both legal substances such as tobacco and alcohol as well as for illicit substances such as marijuana, heroin, amphetamines, and cocaine.

ELEMENTS IN THE DEFINITION OF CRAVING

Although many clinicians and researchers would agree that craving is an important component of addiction, there is considerable debate regarding just what craving means, what causes it, and how to assess it. Because the content and format of a measure of subjective craving will depend, in part, on how one defines the construct, this chapter begins with an overview of several key issues that arise in the definition and conceptualization of craving. One such issue is where to set the threshold separating a mild interest in drinking or drug taking from an urge or desire that is sufficiently intense to be considered craving.

Another question is whether the definition should reflect a combination of emotional, cognitive, behavioral, and physiological experiences that might indicate craving. For example, craving might be experienced as irritation, anxiety, or dysphoria, in addition to or instead of intense desire. Cognitive experiences of craving might include intrusive thoughts and sensory images of the substance, anticipated relief of withdrawal symptoms or negative mood, and expectations of positive outcomes following intoxication. Overt behavioral indications of craving might include rapid or excessive consumption, undertaking arduous or tedious tasks to acquire a drug, selecting drug consumption over money or other valued goods, and difficulty disengaging one's attention from drug-related stimuli. Psychophysiological correlates of craving might include salivation, changes in respiration and blood pressure, and activation of specific brain regions following exposure to drug-related stimuli. Substance abusers often report different combinations of these emotional, cognitive, behavioral, and physiological experiences when they imagine taking drugs or are exposed to drug-related cues.

Another issue when defining craving concerns its time course. For example, craving may be defined and assessed as a relatively acute and short-lived state whose intensity and experiential components shift rapidly depending on both contextual and intrapersonal factors. Craving may also be defined and measured as if it were a more stable (perhaps even trait-like) inclination or propensity to seek drugs somewhat independent of the time since one's last dose, environmental constraints on consumption, and one's overall intention to abstain or continue using drugs. These are not mutually exclusive conceptualizations of craving, and the degree to which one integrates these two views of craving has implications for the content of a questionnaire or behavioral task, the time period over which respondents' experiences will be measured, and the evaluation of a measure's reliability and validity.

Definitions also differ on the degree to which self-awareness is a component of craving. Using a self-report questionnaire to assess craving assumes respondents are aware of and are able to rate the frequency, duration, or intensity of experiences considered indicative of craving. However, some theories of addiction propose that urges and cravings may be elicited without the person's conscious awareness. Unconscious cravings might not be assessable using direct self-report questionnaires but might be revealed using procedures to assess cognitive, behavioral, and physiological experiences considered indirect or proxy measures of craving. Because these and other theoretical considerations have implications for the assessment of craving, the section below provides a short summary of several key etiological explanations of craving.

ETIOLOGICAL EXPLANATIONS OF CRAVING

Classical Conditioning

Craving is often conceptualized as a conditioned response to either environmental or interoceptive stimuli. Over time and with repeated pairings, initially neutral stimuli may become associated with emotional states and withdrawal symptoms that motivate consumption or with the initial effects of intoxication that follow consumption. These conditioned stimuli include locations where alcohol and drugs are consumed, the sight and smell of one's preferred substance, the presence of other drinkers and drug users, and drug-use paraphernalia. After repeated pairings, exposure to these conditioned stimuli may elicit the now conditioned responses of subclinical withdrawal symptoms and/or anticipated intoxication – both of which may be experienced subjectively as craving. Several of the etiological explanations summarized below assume that psychological or psychophysiological processes ultimately resulting in craving begin with exposure to drug-related cues.

Incentive-Sensitization Model

According to this model, both mildly and powerfully intoxicating substances activate "wanting" and "liking" systems in the brain that, in other contexts, are critical to

the survival of humans and other animals. Repeated drug use alters these brain systems such that relevant neural circuits become "hypersensitive" or overreactive to drug stimuli and drug effects. When exposure to drug-related stimuli activates the wanting system, the organism is more likely to desire and seek drugs but may or may not be aware of a subjective experience of craving. Because one implication of this model is that craving may occur without conscious awareness, indirect or implicit measures reflecting attentional bias for drug cues, willingness to work for drug access, or preference for environments associated with past drug use may reflect craving more reliably than self-report.

Obstacles to Automatized Drug Use

According to this model, craving may arise when stimuli trigger automatic drug seeking that is restrained by environmental obstacles to consumption (e.g. one's drug of choice is not readily available) or is restrained by one's commitment to abstain. The physiological and psychological arousal that occur when one exerts effort to overcome environmental obstacles to consumption, or exerts effort to maintain abstinence, may be experienced subjectively as craving. This model implies that craving will be revealed by measures of physiological and psychological arousal and by reports of conscious intentions to consume one's drug of choice. It also implies that craving will occur only when automatic drug seeking is cued in the presence of either environmental obstacles to consumption or when a drug taker has made a commitment to restrain substance use. In the absence of environmental or intrapersonal obstacles, the outcome of automatized drug seeking is not craving, but consumption.

Elaborated Intrusion Theory

Integrating classical conditioning, neuroscience, and cognitive processing perspectives, this theory proposes that a process resulting in craving begins when intrusive thoughts of a target substance are triggered by factors such as physiological changes (e.g. decreasing blood level of the drug), negative affect (e.g. depression), external cues (e.g. sight or smell of preferred drug), and anticipatory responses (e.g. salivation). Most intrusive thoughts are transitory, but when these thoughts are followed by anticipated pleasure or anticipated relief, mental elaboration of the intrusive thoughts may occur. Elaboration combines internal and external information in working memory to create sensory images of the substance and the subjective experience of desire, and one implication of this model is that both these sensory images and subjective desire may be assessed as indications of craving.

Approach-Avoidance Conflict

Arguing that consumption is influenced by more than just an appetite for or desire to take drugs, this conceptualization proposes that craving, consumption, and abstinence are outcomes of inclinations to use a drug interacting with inclinations to avoid using the substance. This model argues that the relative strength or weakness of an addict's inclinations to avoid or restrain consumption impact behavior as much as the strength of one's desires and intentions to take a drug. Therefore, assessment questionnaires based on this model ask about one's attempts to restrain or regulate consumption as well as asking about one's preoccupation with and temptation to consume a drug.

Analogy to Obsessive-Compulsive Disorder

This model is based on the correspondence between the key features of obsessive-compulsive anxiety disorder (OCD) and substance dependence. Specifically, this model proposes that drug craving is analogous to the unwanted, intrusive, disruptive, and anxiety-provoking thoughts or images (i.e. obsessions), and that excessive smoking, drinking, and drug taking are analogous to the repetitively performed and ritualized behaviors (i.e. compulsions), that characterize OCD. Based on this model, questionnaires have been designed to assess the frequency of, duration of, and distress caused by substance-related thoughts, impulses and images and to measure the degree to which substance use and attempting to control one's consumption has interfered with daily functioning.

SELF-REPORTS OF SUBJECTIVE URGES AND CRAVING

The most common method of assessing subjective craving has been to ask respondents to indicate the intensity of their urge or desire for a substance using one or several rating scales. These rating scales – often referred to as visual analog scales (VAS) – are usually anchored by terms such as "none" or "not at all" and "extreme" or "strongest ever experienced," and the respondent is instructed to make a mark on a 100-mm line (or to circle a numeral) to indicate the degree of urge, craving, or desire for the target drug experienced at that moment (or over some specified period of hours or days). Clinicians and researchers may provide a definition of "urge" or "craving" for respondents, but this is not

usually done and respondents are left to interpret the word and indicate their experience of it without further guidance. Paper-and-pencil forms are often used to have respondents record their craving, but electronic devices such as cell phones and handheld computers may also be used to record and transmit one's ratings.

In light of the many advantages of single-item rating scales, such as ease of administration and scoring, suitability for frequent and repeated administration, and apparent sensitivity to rapid changes in craving, such scales are often used in both laboratory and clinical settings. However, single-item rating scales have several disadvantages as well. One concern is that asking the respondent to provide a global rating of the strength or intensity of his or her "urge" or "craving" oversimplifies a complex and perhaps conflicting set of experiences. In addition, the subjective experience of a 10-point increase in craving from 30 to 40 (on a 100-point scale) may feel rather different than the 10-4point increase from 85 to 95 on such a scale. It is also possible that single-item ratings may fail to assess the experience of subjective urges or cravings if the drug taker interprets such feelings as reflecting another psychological or physiological state, such as anxiety, excitement, frustration, ill health, or withdrawal.

Partly to rectify these disadvantages of single-item rating scales, researchers have developed a number of multi-item questionnaires of craving for a wide variety of substances. Adapted from a recent review of such instruments, Table 40.1 provides a summary of these multi-item questionnaires. For each instrument, the table lists the authors of the source document (column 1), the name of the questionnaire (column 2), the target drug and conceptualization of craving on which the instrument is based (column 3), the number and type of items (column 4), the questionnaire's time frame and response options (column 5), and the number and names of factors/subscales, if relevant (column 6). As Table 40.1 shows, these questionnaires reflect one of four basic conceptualizations of addiction and craving: (1) craving as analogous to OCD; (2) craving as one outcome of an approach-avoidance conflict; (3) craving as a multidimensional experience comprised of conscious desire, intentions to consume, anticipated outcomes, and perceived control; and (4) craving as conscious urges measured in terms of their intensity, frequency, and duration. The selection of a questionnaire will be based, in part, on which of these conceptualizations matches the assessor's view of craving.

As Table 40.1 also shows, these questionnaires vary considerably in their length, with the shortest ones containing only four or five items, some containing 12–15 items, and the longest measures containing either 32 items or 45 items in their original versions. Although more items are likely to increase the internal consistency reliability and enhance the content validity of a questionnaire, longer scales have lower utility if the assessor expects craving to be short-lived or conducts repeated measurement of craving within a laboratory or therapeutic session. A shorter questionnaire is also preferred if one is concerned that reading and responding to multiple items might increase existing craving or might elicit craving not otherwise being experienced at that time. Given the advantages of shorter measures, several research teams have revised lengthier initial versions of their questionnaires, and both the longer and shorter versions are summarized in Table 40.1.

Table 40.1 also reveals that these self-report questionnaires differ regarding the time frame over which craving is assessed (e.g. "now," "past week," "when regularly drinking") or do not specify a time frame. Thus, some instruments assess a current state of craving, some assess craving in the recent past, and others measure a more stable preoccupation with or inclination to take a drug. Questionnaires that ask respondents about their craving over some period of hours or days could pose interpretation problems because the assessor does not know whether the drug taker answered items based on an especially salient or peak experience of craving over the designated time period, based on an attempt to average one's craving experiences over that time period, or based on some other rationale. To address these limitations, those using one of these questionnaires might ask respondents about the stability or variability of craving within the designated time period and whether the experience of craving varied depending on context.

The source articles for many of these questionnaires have reported convergent validity among different multi-item questionnaires or between these questionnaires and single-item VAS ratings of craving. These findings suggest that responses on these different instruments reflect some common underlying experience of craving, but the similarity of items across questionnaires and/or respondents' tendency to answer consistently when asked to complete several questionnaires may also account for the correlations among scores on these measures. Although total or subscale scores on different drug-specific questionnaires are often significantly correlated, the coefficients are not so uniformly large that these scales and questionnaires should be considered completely interchangeable.

In addition to evaluating the convergent validity of their measures, some researchers have also evaluated the construct validity of their self-report scales. One type of investigation is based on the assumption that craving scores should increase as blood level of a drug decreases, and several studies have found that craving is more intense following a period of deprivation. Other experiments have shown that asking participants to think about drug-related scenes increases self-reports

TABLE 40.1 Authors and Features of Multi-item Self-report Craving Questionnaires

Authors (Year)	Name of questionnaire	Drug; Conceptualization	Number and type of items	When craving assessed; Response format	Number and names of factors/subscales
Anton, Moak, and Lantham (1995)	Obsessive-compulsive drinking scale	Alcohol; Obsessive-compulsive	14 (with 2 sets of paired items, higher of which contributes to score)	Not specified; 5 choices per item indicating degree of impairment	Two: Obsessive; Compulsive
Franken, Hendriks, and van den Brink (2002)	Obsessive-compulsive drug use scale	Heroin; Obsessive-compulsive	12 (adapted for heroin from OCDS with some items combined or phrasing changed)	Presumably same as OCDS, but not stated in source article	Three: Thoughts and interference; Desire; Control and resistance
Collins and Lapp (1992)	Temptation and restraint inventory	Alcohol; Approach-avoidance	15 (adapted from other questionnaires)	Not specified; 9-point anchored with "never" and "always"	Two: Cognitive and emotional preoccupation; Cognitive and behavioral control
McEvoy, Stritzke, French, Lang, and Ketterman (2004)	Approach and avoidance of alcohol questionnaire	Alcohol; Approach-avoidance	20 items on Australian version; 14 items on American version	Now; Past 24 h; Past week; 9-point anchored with "not at all" to "very strongly"	Three: Inclined to drink; Resolved to avoid alcohol; Obsessed/Compelled
Ikard, Green, and Horn (1969)	Horn–Waingrow scale	Cigarettes; Multidimensional	23 reflecting psychological model of smoking	Not specified; 5 frequency options for each item: always, frequently, occasionally, seldom, never	Six: Habitual smoking; Addictive smoking; Negative affect reduction; Pleasurable relaxation; Stimulation; Sensorimotor manipulation
Leonard, Harwood, and Blane (1988)	Preoccupation with alcohol scale	Alcohol; Multidimensional	14 (items reflect seeking and increased thoughts of alcohol, active drink seeking, heavy drinking)	Not specified; 7-point anchored with "not at all like me" and "very much like me"	One
Tiffany and Drobes (1991)	Questionnaire of smoking urges	Cigarettes; Multidimensional	32 (from pool of 70) with 8 for each of 4 dimensions (desire, positive outcomes, relief from withdrawal/negative affect, intention)	Now; 7-point anchored with "strongly disagree" and "strongly agree"	Two: Intention, desire, and positive outcomes; Relief from withdrawal and negative emotions and strong desire
Cox, Tiffany, and Christen (2001)	Brief Questionnaire of smoking urges	Cigarettes; Multidimensional	10 out of 32 from QSU described above	Now; 100-point scale anchored with "strongly disagree" and "strongly agree"	Two: Strong desire and intention; Anticipation of relief
Tiffany, Singleton, Haertzen, and Henningfield (1993)	Cocaine craving questionnaire	Cocaine; Multidimensional	67 in initial pool → 45 (many adapted from QSU plus new impaired control items)	Now and General (i.e. average over past week); 7-point anchored with "strongly disagree" and "strongly agree"	Four: Desire and intention; Lack of control; Combined positive and negative expectancies; No clear theme

(*Continued*)

TABLE 40.1 Authors and Features of Multi-item Self-report Craving Questionnaires—cont'd

Authors (Year)	Name of questionnaire	Drug; Conceptualization	Number and type of items	When craving assessed; Response format	Number and names of factors/subscales
Sussner et al. (2006)	Cocaine craving questionnaire-brief	Cocaine; Multidimensional	10 (from 45 on the CCQ above)	Now; 7-point anchored with "strongly disagree" and "strongly agree"	One (10 CCQ-Brief items loaded on the same "general craving" factor on full CCQ)
Singleton, Tiffany, and Henningfield (1994; 1995)	Alcohol craving questionnaire	Alcohol; Multidimensional	45–47 (adapted from QSU plus impaired control items)	Now and general; 7-point scale anchored with "strongly disagree" and "strongly agree"	Four: Emotionality; Purposefulness; Compulsivity; Expectancy
Raabe, Grusser, Wessa, Podschus, and Flor (2005)	Alcohol craving questionnaire-revised	Alcohol; Multidimensional	30 items retained from 45 in original ACQ	Now; 7-point scale anchored with "strongly disagree" and "strongly agree"	Two: Urge and intention to drink; Reinforcement
Tiffany, Fields, Singleton, Haertzen, and Henningfield (1993)	Heroin craving questionnaire	Heroin; Multidimensional	45 (adapted from CCQ)	Now; 7-point scale anchored with "strongly disagree" and "strongly agree"	Four: Positive outcomes, intention, craving; Intention and satisfaction; Relief from withdrawal; Lack of control
Ollo et al. (1995)	Questionnaire of cocaine use	Cocaine; Multidimensional	33 (adapted from QSU plus olfactory item)	Now; 7-point scale anchored with "strongly disagree" and "strongly agree"	Four: Anticipation of positive outcomes or relief; Desire and intention to use; Use unpleasant; Uninterested in use
Bohn, Krahn, and Staehler (1995)	Alcohol urge questionnaire	Alcohol; Multidimensional	8 (of 49 in initial pool) loading > 0.70 on primary factor	Now; 7-boxes anchored with "strongly disagree" and "strongly agree"	One (comprised of 4 desire items, 2 positive outcome expectancy items, and 2 impaired control items)
Love, James, and Willner (1998)	Desires for alcohol questionnaire	Alcohol; Multidimensional	36 (based on Clark, unpublished); 14-item short form in same source article	Now; 7-point scale, anchors not provided in source article	Three: Relief from negative states and positive outcomes; Strong desires and intentions; Mild desire and intentions
Heishman, Singleton, and Liquori (2001)	Marijuana craving questionnaire	Marijuana; Multidimensional	47 adapted from previous measures (e.g. ACQ noted above)	Now; 7-point scale anchored with "strongly disagree" and "strongly agree"	Four: Compulsivity or loss of control; Emotionality or anticipated relief negative emotions; Positive outcomes; Immediate intention to use

Heishman et al. (2009)	Marijuana craving questionnaire-short form	Marijuana; Multidimensional	12 (from 47 on the MCQ noted above)	Now; 7-point scale anchored with "strongly disagree" and "strongly agree"	Same four as MCQ: Compulsivity or loss of control; Emotionality or anticipated relief negative emotions; Positive outcomes; Immediate intention to use
Franken, Hendriks, and van den Brink (2002)	Desires for drug questionnaire	Heroin; Multidimensional	36 (adapted from DAQ described above)	Now; 7-point scale, anchors not provided	Three: Desire and intention; Negative reinforcement; Impaired control
Heishman, Singleton, and Moolchan (2003)	Tobacco craving questionnaire	Cigarettes; Multidimensional	45 (32 adapted from QSU, 8 impaired control items; 5 new items)	Now; 7-point scale anchored with "strongly disagree" and "strongly agree"	Four: Relief of negative states; Enjoyment from smoking; Lack of control; Perceived control
Heishman, Singleton, and Pickworth (2008)	Tobacco craving questionnaire-short form	Cigarettes; Multidimensional	12 (3 items from each of the 4 factors in TCQ described above)	Now; 7-point scale anchored with "strongly disagree" and "strongly agree"	Four: Relief of negative states; Enjoyment from smoking; Lack of control; Perceived control
Mol et al. (2003)	Benzodiazepine craving questionnaire	Benzodiazepines; Multidimensional	48 adapted from QSU and CCQ reduced to 32 (deletion of reverse-phrased items) then reduced to 20	Now; 7-point scale anchored with "strongly disagree" and "strongly agree"	One (containing 20 items reflecting 5 dimensions: desire, intention, positive expectancies, relief from negative states, control)
James, Davies, and Willner (2004)	Desire for speed questionnaire	Amphetamines; Multidimensional	40 (adapted from DAQ in Love et al., 1998, plus new items on control of amphetamine use)	Now; 7-point scale anchored with "strongly disagree" and "strongly agree"	Four: Expect positive and negative reinforcement; Strong desires and intentions; Mild desires; Impaired control
Ooteman, Koeter, Vserheul, Schippers, and van den Brink (2006)	Jellinek alcohol craving questionnaire	Alcohol; Multidimensional	24 items from initial pool of 129	Now; While regularly drinking; 5-point anchored with "not at all" and "very much"	One with high inter-correlations among 4 dimension-based subscales (Feelings of urge; Uncontrolled thoughts; Temptation; Physical sensations)
Jimenez, Grana, Montes, and Rubio (2009)	Alcohol craving scale based on three factors	Alcohol; Multidimensional	33 items from initial pool of 40	Over past two weeks of consumption before treatment; 4 options (never, rarely, frequently, always) to indicate frequency with which drinking motivated by each item	Three: Positive reinforcement; Negative reinforcement; Loss of control

(*Continued*)

TABLE 40.1 Authors and Features of Multi-item Self-report Craving Questionnaires—cont'd

Authors (Year)	Name of questionnaire	Drug; Conceptualization	Number and type of items	When craving assessed; Response format	Number and names of factors/subscales
Shiffman and Jarvik (1976)	Shiffman–Jarvik smoking withdrawal questionnaire	Cigarettes; Intensity	6 (of 25 total) that loaded on factor named "craving"	When varies by item (some now, some unspecified); 7-point anchored with "very definitely" and "very definitely not"	One for craving items
Halikas, Kuhn, Crosby, Carlson, and Crea (1991)	Minnesota cocaine craving scale	Cocaine; Intensity, Frequency, Duration	5: Intensity, Frequency/day over past week, Average duration; Change since last week; Effect of medication on craving	Past week; Format varies by item – Intensity: 100 mm VAS; Frequency: 7 choices from none to more-than-20; Duration: 8 choices from 0–5 min to 2 h+	N/A
Voris, Elder, and Sebastian (1991)	Voris cocaine craving scale	Cocaine; Intensity	4: Intense craving-No craving, Depressed-Happy mood, No energy-Too much energy, Feel sick-Feel well	Now; 100 mm VAS for all 4 items	N/A
Weiss, Griffin, and Hufford (1995)	Weiss craving questionnaire	Cocaine; Intensity	5: Current desire; Past 24 h desire; Past 24 h frequency of urge; Past 24 h when cued; Current likelihood if in regular environment	Now; Past 24 h; 9-point with various relevant anchors (e.g. "no desire" and "extremely strong")	One
Flannery, Volpicelli, and Pettinati (1999)	Penn alcohol craving scale	Alcohol; Intensity, Frequency, Duration	5 (frequency/strength/duration of thoughts; ability to resist; overall craving	Past week; 7 item-specific choices reflecting frequency, intensity, duration, difficulty abstaining	One
Kavanagh, May, and Andrade (2009)	Alcohol craving experience questionnaire	Alcohol; Intensity, Frequency, Duration	10 (duration/strength of craving, sensory images, thoughts of drinking)	7 items regarding most recent episode of strong craving; 3 items regarding past 24 h; 5 item-specific choices (e.g. from "weak" or "not at all" to "intense" or "all the time")	One

Adapted from [Rosenberg, 2009. Clinical and laboratory assessment of the subjective experience of drug craving. Clinical Psychology Review, 29, *519–534.]*

of craving on both single-item ratings and multi-item questionnaires. More concrete forms of cue exposure – for example, having drug users look at or handle drug paraphernalia – also elicit reports of elevated craving, usually measured using a single-item self-rating but also occasionally using a multi-item questionnaire.

Research demonstrating the impact of deprivation, drug-related imagery, and exposure to drug photos or paraphernalia on reported craving lends support to the construct validity of self-report questionnaires and VAS ratings. However, despite the practical advantages and theoretical bases for using these three interventions to elicit craving in clinical and laboratory settings, these procedures may not elicit craving or its presumed correlates in every respondent. Furthermore, deprivation and exposure probably elicit a wide variety of emotional and physiological responses in addition to craving, but respondents may attribute such responses to craving rather than other psychological states when craving is the only type of psychological experience about which they are asked. In addition, and probably because it would be difficult to do so, there are no empirical studies evaluating whether cue-elicited craving is subjectively experienced similarly to that which occurs outside the research lab or treatment setting.

As another evaluation of their self-report craving instruments, many investigators have examined the association of single-item and multi-item questionnaire scores with measures of recent consumption, severity of dependence, and number of drug-related problems. One especially innovative investigation compared craving scores as a function of route of drug administration and found that amphetamine abusers who injected the drug reported higher craving than those who smoked or snorted their amphetamines. The association of craving scores with the number of drug-related problems and severity of dependence is often interpreted as lending support to the validity of a questionnaire. However, these associations could also mean that these questionnaires are, in fact, measuring other features of drug addiction (e.g. experience of withdrawal symptoms, negative life consequences) that sometimes covary with, but that many would not consider the same as, the compelling urge or intense desire that defines craving. Depending on one's perspective, associations of self-reported craving with drug consumption and drug-related consequences could be seen as an indication of either good criterion validity or poor discriminant validity.

Limitations of Self-Report Questionnaires

Depending on one's conceptualization of craving, many of these self-report questionnaires may not ask about experiences the clinician, researcher, or drug taker considers indicative of craving. For example, only some questionnaires assess anticipated loss of control, only some include items assessing overt behaviors such as drug seeking, and only several assess bodily sensations that some might consider direct or indirect indications of craving. On the other hand, attempting to represent every thought, emotion, behavior, and physical sensation considered indicative of craving could result in unnecessarily lengthy questionnaires that might be no more reliable and valid than shorter questionnaires.

Self-reports of craving also appear to have limited predictive validity. Specifically, many studies find only a weak correlation between craving measured at a specific point in time (e.g. during intake or at discharge) and one or more aspects of consumption some hours, days, or weeks later. However, other studies report a more robust relationship between craving and subsequent consumption, and further research is needed to evaluate that combination of factors – such as craving, commitment to abstinence, coping ability, and exposure to environmental triggers – that yields the most reliable prediction of posttreatment outcome. We should also recognize that predicting relapse may not be a fair test of any measure of craving given the transitory nature of craving and the multiple influences on consumption in addition to craving.

The validity of craving questionnaires may also be tempered by limitations inherent to almost all self-report questionnaires. These include, for example, intentional and unintentional acquiescence and social desirability biases that might influence a respondent to underreport or overreport one's experience of craving. The quality of self-report measures could also be compromised if respondents limit the amount of effort they exert when reading questionnaire items, when retrieving information about their experiences from memory, and when selecting a response option that best reflects their experience. Craving questionnaires may be especially prone to this type of problem if craving itself disrupts memory and impairs judgment. In addition, assessors and respondents may differ in their understanding or interpretation of questionnaire items, response choices, and rating-scale anchors.

Finally, self-reports of craving might also serve, in at least some cases, as a preconsumption or postconsumption excuse to explain one's purposeful decision to consume a drug. Even if not purposefully misleading, the report of subjective craving may be the story one tells to explain consumption after the fact (e.g. "I must have been craving or I wouldn't have had that first drink"). There has been relatively little research evaluating how these limitations impact the reliability and validity of the instruments listed in Table 40.1, and these questions warrant consideration when one selects – and analyzes the responses from – a self-report questionnaire of craving.

SELF-REPORTS OF OTHER BEHAVIORS CONSIDERED INDIRECT INDICATIONS OF CRAVING

Free-Response Procedure

The free-response or think-aloud procedure allows clinicians and researchers to assess the subjective experience of craving without the use of printed items rated by the participant. The instructions for this procedure are similar to the free association instruction in psychoanalytic therapy because participants are asked to speak their thoughts and feelings aloud without judging how appropriate they might seem; however, in this procedure, substance users free associate during cue exposure or after imagining themselves in a drug-related situation. The participant's verbalizations may be audio- or video-recorded for subsequent content analysis. Advantages of this procedure over printed questionnaires include not having to rely on participants' reading ability or their recall of current or previously experienced craving. Free-response procedures may also reveal unique examples of a respondent's drug-related thoughts and emotions, and video-recording would allow analysis of facial expressions and body posture as subtle signs of craving. However, the think-aloud procedure has its limitations, including the possibility that respondents are unable or unwilling to report every drug-related thought or emotion. There are also practical limitations of recording and transcribing verbalizations for content analysis.

Selection of Drug Access over other Rewards

Designed as a means of assessing the "reinforcement value" of drugs, the multiple-choice procedure (MCP) could be considered an indirect indication of craving. The basic MCP asks respondents to select repeatedly which of two presented options they prefer, typically from a list of choices pairing potential access to a dose of a selected drug against ever-increasing amounts of money (e.g. the choice between a puff on a cigarette or $1.00; next, the choice between a puff on a cigarette or $2.00; next, the choice between a puff on a cigarette or $3.00; etc.). The respondent is informed that, at the conclusion of the testing session, the experimenter will select at random one of the many listed pairs of choices, and the respondent will be given whichever one of the two options (i.e. access to puff on a cigarette or specified amount of money) he/she selected at that choice point. Presumably, the greater one's craving and, therefore, the greater one values possible access to the drug, the more money must be on offer to induce one to switch or crossover from selecting drug access to selecting a specified amount of money. Several studies using this procedure have found higher crossover points when smokers and heroin-dependent patients were drug deprived (and presumably craving more intensely). The MCP holds promise as an easy-to-use and seemingly subtle measure of craving; however, some respondents may always select drug access over money regardless of how weak or strong their craving, and others may select the money option even though they actually prefer the drug. In addition, providing respondents a genuine choice between money and drug access may not be practical or ethical in many clinical and research settings. A more practical though less subtle use of a monetary scale to assess craving has been to ask drug takers how much money they would be willing to spend to obtain a specified dose of their preferred substance.

Drug Dreams

The reported frequency and intensity of drug dreams, and the dreamer's reported reaction to such dreams, could also be considered indirect indications of craving. Depending on one's theoretical orientation, substance-related dreams may be conceptualized as the expression of unconscious wishes to resume drug use, as the processing during sleep of drug cues and craving experienced while awake, or as arousal of neuropsychological systems that underlie both dreams and craving. Although several studies have reported an association between drug dreams and craving while awake, the prognostic value of such dreams as a predictor of lapse or relapse has been mixed. Furthermore, research has yet to assess whether the content of such dreams (e.g. searching for drugs, preparing to consume, refusal of drug offers) and the dreamer's reaction to the dream content (e.g. anxiety, guilt, subjective intoxication, relief that "it was only a dream") reliably precede or follow other indications of craving, drug seeking, and drug use while awake. The value of reports of substance-related dreams as an indirect measure of craving would benefit from further research to assess whether the frequency and content of drug dreams are influenced by presleep drug deprivation and presleep cue exposure designed to induce craving.

OBSERVATION OF BEHAVIORAL AND PHYSIOLOGICAL INDICATORS OF CRAVING

Motoric Behavior

Craving could also be defined in terms of overt behaviors, such as willingness to work for drug access, latency to consume given access, preference for drug-use

environments over other locations, and speed of consumption. Although assessing these behaviors obviates the need to rely on potentially misleading self-report of an internal state of craving, measuring craving only in terms of overt behavior raises other problems. For example, even intense craving does not always result in smoking, drinking, or drug use. Furthermore, both recreational and pathological drug use are influenced by a variety of factors other than craving, including mood, personal consumption goals, availability of the drug, and anticipated outcomes of consumption. In short, the overt behaviors noted above are influenced by more than just craving.

Attentional Bias for Drug Cues

Attention and attentional bias facilitate mental processing of the numerous internal and environmental stimuli to which we are exposed in our everyday lives. There is considerable debate regarding how and why attentional bias occurs among substance abusers, and mixed evidence regarding the reliability and validity of attentional bias for drug-related stimuli as an indication of addiction and predictor of consumption. Nonetheless, attentional bias has been proposed as both a cause and outcome of craving, and performance on attention-related reaction-time tasks such as the addiction-Stroop and dot-probe procedures may serve as proxy measures of subjective craving.

In the addiction-Stroop procedure, drug-related words and "neutral" or contrast words are used as stimuli, and the respondent is asked to name the colors of the ink in which the words are printed (or in which stimulus photos are tinted). A recent meta-analytic review of research using the addiction-Stroop concluded that longer reaction times on this task often distinguish addiction-related words from contrast words, distinguish substance abusers from nonabusers, and predict relapse following a period of abstinence. However, it is not clear how much slower one's reaction time to drug words (vs. contrast words) must be to indicate craving. In addition, performance on the addiction-Stroop could be influenced by several types of cognitive processes in addition to or instead of craving (e.g. reading drug-related words might elicit disruptive anxiety or might activate drug-related memories that distract one from the color-naming task).

Another reaction-time task, the dot-probe procedure, may also be considered a proxy measure of craving. On each of the many trials comprising this task, a drug-related word or picture and a neutral/control word or picture are displayed simultaneously on a computer screen, typically for just a fraction of a second. To control for handedness and other biases, the left–right position of the photos is counterbalanced across trials. After the pair of images is removed from the screen, a probe stimulus such as a dot, arrow, or asterisk appears in the location occupied previously by one or the other of the stimulus pictures. Participants are asked to indicate the position of the probe by pressing one of two response keys as quickly as possible. Attentional bias (and perhaps craving) may be inferred if reaction times are faster when the probe replaces drug-related stimuli than when it replaces neutral stimuli, revealing that one's attention has been "captured" and remains focused on the same side of the screen recently occupied by the drug-related stimulus. Consistent with this hypothesis, a recently published meta-analysis found a significant, though relatively weak, association between subjective craving and various measures of attentional bias (including the dot-probe). Once again, however, the point at which performance on the dot-probe task reflects an urge or craving remains to be delineated. Perhaps attentional bias must remain focused on drug stimuli for at least several seconds for craving to be elicited or intensified, or perhaps only when craving is of sufficient intensity might it result in scanning the environment for and attentional bias to drug-related cues.

Psychophysiological Reactivity

There are several rationales for assessing psychophysiological responses – such as salivation, changes in skin temperature, respiration or blood pressure, increased perspiration, and activation of specific brain regions – as indirect measures of craving. First, self-awareness of subjective craving could itself elicit psychophysiological arousal; secondly, conditioned stimuli previously associated with drug withdrawal or drug consumption could elicit both subjective craving and psychophysiological arousal simultaneously; thirdly, psychophysiological responses of which the person is not consciously aware could serve as conditioned stimuli that elicit craving; and fourthly, self-awareness of changes in physiological arousal could be interpreted by the drug user as subjective craving (e.g. "My palms are sweating, and so I must really need a drink").

Consistent with these rationales, research has found that exposure to drug-related stimuli, compared to neutral/comparison stimuli, often yields changes in measures of both peripheral and cortical functioning. Although physiological reactivity could provide an apparently "objective" measure of either conscious or nonconscious craving, psychophysiological reactivity – like craving itself – is a multi-determined phenomenon influenced by such factors as individual differences in

baseline reactivity and the environmental context in which reactivity is measured. Furthermore, physiological measures have limited clinical applicability given the need for special training to operate recording equipment, safety and esthetic concerns if one collects bodily fluids such as saliva, and the sometimes weak correlation between physiological reactivity, cue exposure, and other presumed measures of craving.

CONCLUSION

Developing and evaluating measures of subjective craving are complicated by questions regarding the emotional, cognitive, behavioral, and physiological components of craving, the threshold separating preference from craving, the degree to which one is interested in an acute and fluctuating experience of craving or a relatively stable preoccupation with or inclination to use a target drug, and the degree to which substance users are aware of and able to report on their motivational state of craving, among others. These questions notwithstanding, researchers and clinicians have a variety of assessment methods from which to choose, including single-item ratings of craving, multi-item questionnaires representing several different conceptualizations of craving, the free-response or think-aloud procedure, reports of drug dreams, observation of drug seeking and drug consumption, performance on reaction time tasks, and physiological reactivity.

Given the diversity of procedures for assessing craving, on what basis might clinicians and researchers choose a specific measure? Each type of assessment procedure has its advantages and disadvantages, and most could be used to measure craving experienced during a period of abstinence, within an ongoing drug-use episode, or during withdrawal. Selection will depend on several factors, including one's theoretical conceptualization of addiction and craving, the time frame over which one wants to assess craving (e.g. right now; over past 24 h; over past week), access to physiological recording equipment, opportunities to observe drug consumption, and the need for quick assessment when conducting repeated measurement or if respondents have limited concentration. Depending on one's setting, the expensive equipment needed for some procedures (e.g. assessment of brain activation), and the ethical and practical limitations of procedures that involve providing access to an illicit drug, will also influence one's selection. Perhaps the best strategy, if it is practical in one's clinical or research setting, is to employ several measures, each assessing different emotional states, cognitions, and behaviors considered indicative of craving.

Finally, as the concept of addiction has been applied beyond substances to apparently "addictive" activities such as gambling, physical exercise, computer use, and viewing sexually explicit material, there have been suggestions that researchers and clinicians also assess urges or cravings for those activities. For example, several of the multi-item self-report instruments listed in Table 40.1 have been modified to assess urges or inclinations to gamble. Further research may reveal that activity-specific modifications of craving questionnaires, and the other self-report, behavioral and physiological procedures outlined above, will improve the assessment and treatment of other types of addictive disorders.

SEE ALSO

Deprivation, Craving, and Affect: Intersecting Constructs in Addiction

Glossary

Addiction-Stroop procedure task during which participants are asked to read a list of drug-related words and "neutral" or contrast words, and to name the different colors of the ink in which the words are printed. Longer reaction times to drug-related versus contrast words may be interpreted as attentional bias.

Attentional bias reflexive, automatic awareness and cognitive processing of emotionally important stimuli in one's environment.

Craving subjective experience of an intense or compelling urge, desire or inclination to consume a drug.

Cue exposure procedure during which drug takers are shown drug-related photos, videos, or actual paraphernalia to assess cognitive, emotional, or physiological reactions considered indicative of craving (or to provide practice coping with craving).

Dot-probe procedure task involving multiple simultaneous presentations of both a drug-related word or picture and a "neutral" or contrast word or picture, typically for a fraction of a second. Following offset of the stimuli, a probe stimulus such as a dot, arrow, or asterisk appears in the location occupied previously by one or the other of the stimuli. Respondents are asked to indicate the position of the probe by pressing one of two response keys as quickly as possible. Shorter reaction times across those occasions when the drug stimulus and dot appear on the same side of the screen are interpreted as attentional bias.

Drug dreams dreams during which a person imagines seeking, consuming, or refusing drugs; such dreams may be an indirect indication of craving.

Free-response/think-aloud procedure respondents are asked to speak their thoughts and feelings aloud during cue exposure or after imagining themselves in a drug-related situation.

Multi-item craving questionnaire self-report questionnaire containing at least several and often dozens of items asking about various emotional, cognitive, and behavioral experiences considered indicative of craving for a specified substance.

OCD obsessive-compulsive anxiety disorder.

Visual analog scale (VAS) 100-mm line on which the respondent makes a mark to indicate the degree of urge, craving, or desire for the target drug experienced at that moment (or over some specified time period).

Further Reading

Berridge, K., Robinson, T., 1995. The mind of an addicted brain: neural sensitization of wanting versus liking. Psychological Science 4, 71–76.

Drummond, D.C., 2001. Theories of drug craving, ancient and modern. Addiction 96, 33–46.

Ferguson, S.G., Shiffman, S., 2009. The relevance and treatment of cue-induced cravings in tobacco dependence. Journal of Substance Abuse Treatment 36, 235–243.

Field, M., Munafo, M.R., Franken, I.H.A., 2009. A meta-analytic investigation of the relationship between attentional bias and subjective craving in substance abuse. Psychological Bulletin 135, 589–607.

Franken, I.H.A., 2003. Drug craving and addiction: integrating psychological and neuropsychopharmacological approaches. Progress in Neuro-Psychopharmacology & Biological Psychiatry 27, 563–579.

May, J., Andrade, J., Panabokke, N., Kavanagh, D., 2004. Images of desire: cognitive models of craving. Memory 12, 447–461.

Pavlick, M., Hoffmann, E., Rosenberg, H., 2008. A nationwide survey of American alcohol and drug craving assessment and treatment practices. Addiction Research and Theory 16, 591–600.

Rosenberg, H., 2009. Clinical and laboratory assessment of the subjective experience of drug craving. Clinical Psychology Review 29, 519–534.

Sayette, M.A., Shiffman, S., Tiffany, S.T., et al., 2000. The measurement of drug craving. Addiction 95 (Suppl. 2), S189–S210.

Stritzke, W.G.K., McEvoy, P.M., Wheat, L.R., Dyer, K.R., French, D.J., 2007. The Yin and Yang of indulgence and restraint: the ambivalence model of craving. In: O'Neal, P.W. (Ed.), Motivation of Health Behavior, pp. 31–47. Nova Science Publications, Hauppauge, NY.

Tiffany, S.T., 1992. A critique of contemporary urge and craving research: methodological, psychometric, and theoretical issues. Advances in Behavior Research and Therapy 14, 129–139.

Tracy, J.I., 1994. Assessing the relationship between craving and relapse. Drug and Alcohol Review 13, 71–77.

Wiers, R.W., Stacy, A.W., 2006. Implicit cognition and addiction. Current Directions in Psychological Science 15, 292–296.

CHAPTER

41

Deprivation, Craving, and Affect: Intersecting Constructs in Addiction

Jason A. Oliver, David A. MacQueen, David J. Drobes

Moffitt Cancer Center, University of South Florida, Tampa, FL, USA

OUTLINE

INTRODUCTION

Substance dependence and addictive behavior are complex syndromes, driven by a number of different processes. However, three concepts that are thought to play a central role in nearly every theory of addiction are deprivation/withdrawal, craving, and affect. The experience of drug craving and affective disturbance are key features of drug dependence that are acknowledged by nearly every theory of addiction. Most notably, both symptoms of dependence are markedly pronounced during drug deprivation and are thus characteristics of withdrawal. Though each of these concepts is distinct, they are also inextricably linked and interact in a number of ways. In this chapter, we will review the current state of both psychosocial and neurophysiological evidence derived from both nonhuman animal and human research on these topics. We provide basic definitions and formulations of these topics individually, and then discuss both specific research findings and overarching models that attempt to integrate these concepts.

DEPRIVATION/WITHDRAWAL

Individuals who are deprived of a drug after an extended period of regular use frequently experience a number of symptoms, with the cluster of symptoms being referred to as withdrawal. Withdrawal from substances can include changes in appetite and weight,

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00041-3

nausea, sleep disturbance, sweating, aches and pains, diarrhea, seizures, convulsions, and more. Symptoms can range in severity from minor irritants that may not be perceptible to the individual experiencing them, to serious conditions that carry a risk of permanent injury or death. These symptoms can emerge as early as 30 min to a few hours after the time of last drug use and can persist for days, weeks, or even months. Not all dependent users will experience significant withdrawal symptoms when usage is stopped, and the symptoms they do experience may change over time within a single withdrawal period (i.e. new symptoms appearing 1 week into the deprivation period as others are remitting), or across episodes of withdrawal (i.e. different symptoms experienced from one attempt to stop drug use to another). Different symptoms may follow different time courses during withdrawal, with some symptoms emerging later, or lasting longer. During an episode of withdrawal, administration of the drug to which the individual has become dependent will alleviate the withdrawal syndrome. Although the specific symptoms of withdrawal vary substantially across drug types, and span a variety of domains (e.g. physical, cognitive, etc.), two withdrawal symptoms that occur almost universally across drug types are craving and disturbance in affect.

Tolerance and Withdrawal

A discussion of drug withdrawal requires some understanding of the mechanisms of chemical dependence. Repeated administration of drugs of abuse often results in the development of tolerance, in which the potency of the drug (the amount of the substance required to produce an effect) is reduced across time. Tolerance can be described as a reduction in the physiological or subjective effects of a drug across repeated administrations of the same dose, or as a pattern in which increasing amounts of the substance are needed to achieve the same effects produced previously by lower doses. The mechanisms underlying tolerance can be described as dispositional, functional, or learned/behavioral. Both dispositional and functional tolerances result from physiological responses to the maintained presence of a drug within the system. Dispositional tolerance results from processes that reduce the distribution of the drug to its site of action, such as increased production of enzymes that break down the substance before it can exert its effects. Functional tolerance results from physiological changes that cause sites of action to be less responsive to the presence of the drug. As a direct result of the development of functional and dispositional tolerance, drug abstinence after a period of maintained exposure often results in the production of physiological states opposite of those produced by the drug. This abstinence effect is the withdrawal syndrome referenced above and, as previously noted, is associated with reports of negative affect and drug craving in humans, as well as drug-seeking behaviors in both human and animal models of drug dependence.

CRAVING

Despite the proliferation of research on craving in recent decades, scientists have yet to agree upon a precise definition of craving. It is generally agreed that it is an approach-oriented state that reflects the desire or urge to consume, though there is substantial disagreement over how strong a desire must be to constitute craving. While some view craving and urge as interchangeable terms that exist along a continuum, others reserve the term craving solely for describing extremely strong urges. Craving is generally subdivided into two categories. Tonic craving reflects a generalized state of craving that persists over long periods of time, though potentially fading into the background during periods of active substance use. It is typically linked to an individual's dependence level and is exacerbated by withdrawal. As such, tonic craving can be described in the context of physiological tolerance (dispositional and functional). As physiological tolerance develops, deviations in active levels of the drug present in the body that occur during the course of repeated administration can have more profound withdrawal effects, including increases in tonic craving.

In contrast, phasic (or, cue-elicited) craving is evoked by particular internal or external stimuli, such as the presence of a particular stimulus that is closely associated with prior drug use, a stressor, or even a drug-related thought. Importantly, these two forms of craving appear to be additive. That is, reports of craving are likely to be strongest in situations when cues are present, and withdrawal is being experienced, relative to either situation occurring in isolation. Cue-elicited cravings are typically thought to be relatively short-lived but can be experienced quite intensely. Despite this intensity, it is important to note that while there is substantial evidence linking tonic craving to actual drug use and risk for relapse, phasic craving has produced mixed evidence. While it may seem that craving would be closely linked to actual drug use behavior, and risk for relapse following an attempt to stop drug use, the actual link between craving and drug use appears to be far more complicated. Relationships between craving and drug use are small, or even absent, in many instances of cue-elicited craving.

Recently, scientists have also begun to distinguish between proximal cues, those which are directly related to substance use, such as drug-related paraphernalia or the drug itself, and distal cues, which may be highly individualized and will not necessarily have obvious

ties to substance use for all individuals. Distal cues might be a certain location where a person frequently engages in drug use (e.g. a chair one sits in while drinking), or an object or activity one typically is involved with while substances are being used (e.g. a phone one talks on while smoking).

Physiological Processes and Craving

In addition to self-reported craving, exposure to drug-related environmental stimuli can also be associated with discrete neural, somatic, or autonomic activity, even though no substance has been administered. These physiological responses may reflect the activation of motivational states, or other mechanisms that support drug-seeking behavior. Because cue-elicited cravings are produced by exposure to drug-related stimuli and environments rather than by withdrawal, they are able to produce experiences of craving even after long periods of drug abstinence. As a result, phasic craving produced by learned/behavioral processes are often separated from cravings occurring as a result of active dependence and acute withdrawal (tonic craving). There is evidence that both forms of craving share common neural substrates. Activity in the mesolimbic dopaminergic pathway, consisting of dopaminergic neurons in the ventral tegmental area (VTA) that project upon nucleus accumbens (NA), has been associated with both forms of craving. Because increased activation of this pathway has been consistently associated with the reinforcement (increased frequency) of drug-taking behaviors, these regions have also been assumed to underlie the pleasurable experience of drug use and are often referred to as the "pleasure pathway." It has been suggested that the subjective pleasure derived from drug use can be separated from the experience of drug wanting (craving). This is supported by the observation that repeated administration of abused substances frequently leads to progressive reductions in the self-reported pleasurable effects of the drug (tolerance), while self-reported craving increases across time. Thus, the effects of chronic drug administration on craving may be best described as sensitization, in that repeated delivery of a consistent dose leads to increases in craving both during withdrawal and in response to environmental stimuli.

AFFECT

Definitions

Another withdrawal symptom that is pervasive across drug types is disturbance in affect. Though precise definitions of affect remain a subject of controversy, most scientists now distinguish affect from the related concepts of mood and emotion. Emotion generally refers to brief states that manifest across a variety of component systems (e.g. subjective, behavioral, physiological). These are typically intense despite their transience, are thought to last no more than a few minutes, and frequently occur in response to a discrete, identifiable stimulus. In contrast, moods persist for longer durations (e.g. hours, days, or even weeks), vary more widely in intensity, and do not necessarily manifest across as many component systems as emotions do. Experiencing intense emotional states can have significant physical consequences, such as a substantial increase in heart rate, and therefore are not typically sustainable for extended periods of time. Fortunately, while individuals spend the majority of their existence in one mood state or another, emotions occur less frequently. Affect refers to an overarching construct that encompasses both mood and emotion. Scientists also generally distinguish between short-term changes in affect (i.e. state affect) and the relatively stable individual differences in affect that may persist across years.

Models of Affect

Although research has had some success with defining core affective states, work examining the overall structure of affect has generally supported a dimensional view. That is, rather than moods occurring as discrete, easily separable states, they tend to be highly correlated. Although individuals may identify specific moods (e.g. fear, sadness), these are frequently associated with other negative states, such as guilt, anger, and anxiety. Consequently, models positing a two-dimensional structure of affect have dominated the field. One of these models was developed in an attempt to disentangle the valence component of affect (e.g. happy vs. sad) from the arousal component (e.g. active vs. drowsy). A related model has also been developed, which retains this distinction between the valence and arousal components of affect, but also allows for the simultaneous experience of both positive and negative affects. Thus, an individual can experience both anxiety and excitement surrounding an upcoming event. Finally, though it is not frequently acknowledged in the scientific literature, it should be noted that the distinctions between emotion and mood have clear parallels to those between cue-elicited and tonic forms of craving.

NEGATIVE REINFORCEMENT

Development of Negative Reinforcement Effects

While escalating drug use after initiation is often driven by the direct euphoric effects produced by

drug taking, these euphoric effects generally lessen across time as physiological (functional and dispositional) and learned tolerance develop. As a result of these homeostatic processes, deprivation after the development of tolerance can induce negative states (withdrawal) that are alleviated only by the administration of the same or a related drug. Thus, early models of addiction posited that across time, the motivation for drug use shifts from sensation seeking (euphoria) to alleviation of withdrawal. According to these models, rather than eliciting euphoric effects, drugs were used by dependent individuals primarily to return to baseline mood following increased levels of negative affect due to withdrawal. Given that negative affect symptoms can emerge very quickly after drug use is stopped, affect may be a critical element of withdrawal that motivates continued drug use. The pattern of withdrawal symptoms may also play a role in how distressing they are. The immediacy of certain withdrawal symptoms, such as negative affect, may make them more potent motivators of drug-use behavior than other symptoms that develop slowly across a period of deprivation.

Addiction and Negative Reinforcement

Negative reinforcement models have dominated contemporary theories of addiction. These models have emphasized the role of withdrawal relief or anticipation of withdrawal relief in both craving and drug use. According to these models, drug-seeking and drug use behaviors are rewarded by the removal or avoidance of aversive states associated with withdrawal, which subsequently leads to craving in future situations in which these symptoms are anticipated or experienced. The subjective discomfort of withdrawal may be partially determined by a number of nondrug factors, including differences in genetic makeup, personality, and environment. One area that has received increased attention in recent years is the effects of menstrual cycle phases on withdrawal symptoms for female substance users. Though preliminary, there is evidence that withdrawal symptoms may worsen during the luteal phase of the menstrual cycle. However, disentangling withdrawal symptoms from premenstrual symptoms has proven a significant challenge and further research will be needed before definitive conclusions can be drawn.

Environmental Influences on Negative Reinforcement

As mentioned previously, environmental stimuli that have been associated with drug use can induce drug cravings and drug seeking in humans, as well as physiological effects. Similarly, specific stimuli and environments associated with drug administration can produce drug-seeking behaviors in animals. Evidence from recent animal studies suggests that environmental cues may actually produce stronger cue-elicited drug seeking (sensitization) across lengthier periods of abstinence. For instance, rats trained to self-administer cocaine and then withdrawn from the drug show more drug seeking (production of the trained self-administration behavior) after being returned to the drug administration context the longer the animal has been deprived of the drug. The enhanced effects of prolonged deprivation from abused substances on the production of drug-seeking behavior in response to environmental stimuli exceeds deprivation effects observed with other natural reinforcers (such as sugar) on the later production of trained behaviors in response to environmental stimuli. These observations of increased cue-elicited drug seeking as a function of deprivation period (referred to as incubation) have been associated with increased levels of brain-derived neurotrophic factor (BDNF) – a growth factor that enhances nerve cell growth and survival. Specifically, levels of BDNF have been shown to increase across periods of extended withdrawal in the NA, the VTA, and the amygdala – an effect not seen for other natural reinforcers.

While research on craving incubation in humans remains in its infancy, preliminary evidence suggests that cigarette cravings may incubate across periods of withdrawal. Should the phenomenon be consistently documented in humans, animal models of incubation will no doubt be critical in evaluating the effects of novel compounds that prevent, reverse, or otherwise alter the neuroadaptations underlying the sensitization of drug craving, including longlasting changes in responsiveness of the mesolimbic dopaminergic pathway. Across numerous drugs of abuse, imaging studies in dependent substance users suggest that chronic substance use results in reductions in dopamine release and in dopamine receptor density. When compared with nondependent controls, drug-dependent subjects demonstrate decreased responsiveness of the mesolimbic dopaminergic pathway to both drug administration and other natural rewards. This effect is congruent with the observation that substance-dependent individuals frequently report reduced enjoyment of nondrug-related activities. Anhedonia – an inability to derive pleasure from activities – is a characteristic symptom of depression. Thus, drug-induced changes to intrinsic reward pathways may serve to exacerbate or produce this facet of depression across long-term use.

Negative Reinforcement and Contemporary Theory

One prominent negative reinforcement theory of drug use emphasizes that drug use may be driven by affective withdrawal symptoms that can occur outside of conscious awareness. That is, as withdrawal symptoms begin to develop, an individual may take drugs to avoid experiencing those negative effects even before becoming fully aware that they were emerging. Over time, the addicted individual may become conditioned to expect a reduction in negative affect as a result of drug use because the sensation of relief from negative affect experienced as a result of withdrawal is generalized to other instances of negative affect. Thus, a regular drug user may frequently experience decreases in negative affect as a result of drug use, but this occurs only due to relief of withdrawal symptoms that emerged as a result of regular drug use. However, the individual may believe that drug use is capable of relieving negative affect in other distressing situations independent of withdrawal. In turn, craving can occur as the individual becomes highly motivated to use drugs in order to escape or avoid the experience of negative affect.

SELF-MEDICATION MODELS

One significant challenge for scientists examining negative reinforcement models of addiction has been distinguishing the transient effects of drug withdrawal from offset effects. While true withdrawal symptoms are caused directly by physiological adjustment to the presence (and subsequent absence) of the drug, they may be confounded by underlying, trait-like conditions that were masked by continued drug use. Unlike withdrawal symptoms, these trait effects will not necessarily diminish over time. While not all individuals who become addicted to a drug of abuse suffer from mood and anxiety disorders, rates of drug use are substantially higher among populations with psychiatric illness (including depression) than among individuals without any diagnoses. Moreover, individuals prone to affective disturbances are more likely to experience depression as a result of withdrawal from substances and tend to report greater withdrawal symptoms overall. These individuals may be more highly motivated to use drugs than other drug users, experiencing more frequent and more severe urges. Relatedly, they may have difficulty differentiating withdrawal symptoms from more stable, individual differences. For example, an individual with recurrent major depressive disorder experiencing the onset of a depressive episode may misattribute this decrease in mood to drug withdrawal. As a result, the individual may be more motivated to engage in drug use behavior and their dependence on the drug may increase. Although negative affect is a symptom of withdrawal from all major classes of drugs, misattributions that are more specific to particular drug classes may also play a role in driving drug use behavior. For example, an opioid addict may be unable to differentiate the flu-like aches of opioid withdrawal from an underlying chronic pain condition, and a benzodiazepine addict with chronic insomnia may be unable to differentiate withdrawal-induced insomnia from a return to baseline sleeping difficulties. Of course, functioning can vary widely even in nonclinical samples, and further research will be needed to determine the extent to which individual differences such as personality, genes, and other factors may cause subsets of individuals to experience differential effects resulting from substance use and deprivation.

Interactions of Affect and Cognition

Studies of how mood motivates drug use have not typically examined individual differences in baseline mood. Thus, it is possible that a subset of individuals, such as those with an active mood disorder, would experience a differential benefit of the mood-altering effects of abused substances. Self-medication of cognitive deficits may also occur, as thought to be the case for individuals with a variety of disorders that impair cognition (e.g. Attention Deficit Hyperactivity Disorder). It is important to note that at least in the case of alcohol and nicotine, the effects of the substance on affective state may be mediated by the cognitive effects of the substance. One such study revealed that smoking a cigarette did indeed relieve experimentally induced negative affect, but only when the smokers were given something else to focus their attention on after smoking. Similarly, studies have shown that alcohol may reduce anxiety by decreasing attention to threat cues when distracting stimuli are present.

POSITIVE REINFORCEMENT

Role of Positive Reinforcement in Drug Use

While research on addiction has traditionally emphasized the role of negative reinforcement (e.g. withdrawal relief) in motivating drug use, this contrasts sharply with the well-documented euphoric effects that exist across the majority of substances of abuse. Addiction must progress and tolerance must typically develop in order for significant withdrawal symptoms to emerge, so negative reinforcement models struggle to explain how drug use is motivated early in the course of addiction, before use has escalated. Similarly, negative reinforcement

does not account for the subjective experience of drug addicts, who frequently report drug use to enhance pleasant situations (e.g. parties). A history of euphoric effects as a result of drug use would help explain this association. Research across both animals and humans indicates that drugs of abuse can function similarly to other positive reinforcers, such as food. However, while most substances of abuse have verifiable euphoric effects following use (e.g. cocaine, opiates), this is not true in all cases. Although tobacco smokers subjectively report smoking to enhance positive affect, there is scant evidence to support the idea that nicotine is actually effective in doing so. Rather, the mood-altering effects of nicotine appear to be driven primarily by withdrawal relief, even among lighter smokers. It has also been suggested that the decreased responsiveness to natural rewards (as a result of the development of tolerance within reward pathways) may serve to increase the comparative saliency of direct drug effects, which serve to increase dopamine levels in the mesolimbic dopaminergic pathway. That is, drug use may dampen responses to other rewards, causing a relative increase in the subjective pleasure associated with drug use relative to these other rewards. While chronic drug use generally produces tolerance to drug effects on dopamine release within reward pathways, the administration of drugs (especially with increasing doses) may allow for dopaminergic activity that is comparatively greater than what can be achieved by other natural reinforcers. As a result, chronic substance users can develop a preference for drugs over other rewarding activities. The development of functional tolerance within reward pathways might represent a mechanism through which drug craving/seeking becomes sensitized. Craving may in fact be driven by this sensitization to drug-specific reward, as the pleasurable feeling obtained from the drug may exceed that of alternatives. This hypothesis suggests that drug cravings assessed during periods of dependency and craving during immediate withdrawal will be related to the development of tolerance within reward structures. Thus, the frequency, active dose, and duration of chronic administration will have a direct impact on drug craving during dependence and immediate withdrawal (tonic craving).

Drug Sensitization

If craving is indeed related to neuroadaptations occurring within reward pathways that are consistent across drugs of abuse, then it is also expected that adaptations occurring as a result of administration of one substance should facilitate the development of craving for a second substance. This prediction is supported by evidence from animal studies in which drug-seeking behavior (such as self-administration) is facilitated by prior exposure to drugs. It has been demonstrated in several species that prior administration of a drug can produce enhanced self-administration of the same drug. In addition, a number of studies have now reported cross-sensitization in which the prior administration of one drug enhances measures of craving/drug seeking for another drug with a different mechanism of action. For example, rats pretreated with amphetamine have enhanced preference for contexts associated with the delivery of morphine and vice versa, even though the two drugs act through very disparate mechanisms and have differing profiles of behavioral effects. In animal studies drug craving is inferred by the extent an animal will produce trained behaviors to receive a drug or the animal's preference for contexts associated with drug effects. The implication drawn from this research is that the development of craving for one drug in humans may serve to enhance the development of craving for other abused substances with the same or with different behavioral effects. Such effects have been observed in human research: drugs with euphoric effects can also increase both craving and consumption of other drugs, and these effects may be partially mediated by the original drug's effects on mood. Similarly, presentation of a cue for a specific drug can induce craving for other drugs that are used contemporaneously with the presented drug. Although this effect likely emerges as a result of classical conditioning, cross-drug priming may play a role in establishing the patterns of contemporaneous use that allow this conditioning to occur. Given the frequency with which individuals abuse multiple substances, cross-drug effects may play an important role in the development of secondary addictions and escalation in overall substance use.

EXPECTANCY THEORY

Although researches on the reinforcing aspects of substance use have generally emphasized direct drug effects on relevant motivational processes (i.e. escape/avoidance of aversive states or enhancement of positive states as causes of craving and drug-seeking behavior), they generally fail to explain the initiation of substance use. Beliefs about the effects or outcomes of drug use can develop long before initiation of use. In addition, these beliefs may continue to evolve after initiation, even independent of actual drug effects. That is, people may anticipate certain drug effects and this expectancy may drive behavior even when these expectations are not realistic. In other words, even if drug use may not actually result in the intended effects (i.e., relief of negative affect), the belief that the substance will have this effect might result in craving and motivate drug-seeking

behavior. At least in the case of nicotine, there is evidence that nicotine does not effectively relieve negative affect – except in the case of nicotine withdrawal. Although context (e.g. presence of distractors) may play an important role in moderating this, individuals may generalize their experiences of nicotine relieving negative affect due to withdrawal to other situations involving negative affect. Both personal experiences with a drug, and information gained from other sources of information (e.g. peer users, marketing) play a role in shaping expectancies. Indeed, these expectancies begin to develop at a fairly young age and continue to develop over time. Research indicates that negative expectancies decrease and positive expectancies increase as children enter adolescence, the time when substances are frequently used for the first time. The fact that expected effects of these substances are malleable also makes them a potential target for treatment. There is evidence that reducing the positive expectancies related to drinking can lead to decreased consumption.

AFFECT-INDUCED CRAVING

Previously, withdrawal was discussed as a causal influence that affected both craving and affect. However, it has long been recognized that changes in mood or emotional state can serve as powerful triggers for drug craving independent of withdrawal. Extensive laboratory research has documented that induction of negative affective states, even in the absence of direct substance use cues, can reliably induce drug cravings across a variety of substances (i.e. smoking, alcohol, cocaine). Similarly, stressful or otherwise unpleasant events can produce a high risk for relapse. Research with alcohol indicates that affect may interact with discreet substance use cues to predict drug use behavior above and beyond the level of prediction offered by either cue in isolation. In contrast, there is mixed evidence to support an increase in craving induced by positive affective states. Some research has shown that risk for relapse is also high during celebratory or other pleasant events. However, it has been suggested that this effect is driven by the increased presence of drug-related stimuli in pleasant situations (e.g. alcohol at a party) rather than the positive mood itself. Nonetheless, there is substantial evidence that drugs with potent ability to increase positive mood (e.g. cocaine, alcohol) can increase the likelihood of ingesting other drugs.

Negative Affect, Craving, and Physiology

Because negative affect has long been recognized as a precipitant of craving and drug relapse in humans, attempts have been made to model this phenomenon in animals. Though it is difficult to conceptualize or measure negative affect in nonhumans, animal models of the interaction between stress and drug administration have been fruitful in uncovering the neural mechanisms involved. Such studies have been conducted in a similar manner to those investigating the role of environmental stimuli in drug craving. Animals are trained to produce a behavior (e.g. lever pressing) to receive doses of an abused substance. Subsequently, the animal is deprived of the substance by providing no drug administration after production of the trained behavior. Eventually, animals cease producing the behavior. After a period of time a subgroup of the deprived subjects then receive a stressful stimuli such as a foot shock or a tail pinch. Subjects receiving stressful stimuli tend to begin producing the behavior that previously resulted in drug delivery, while subjects not receiving stress do not. This is referred to as reinstatement, and it has been demonstrated that stressful stimuli can reinstate responding for a range of substances including cocaine, nicotine, alcohol, heroin, and amphetamines. Reinstatement of drug-seeking behavior after stressful stimuli has been associated with increases in stress hormones (corticotrophin releasing factor, CRF) and noradrenaline in the amygdala, a structure associated with fear, negative affect, and emotional processing of memory in humans. In addition, increased levels of CRF, glutamate, and dopamine are observed in the VTA. Selective inactivation of the portions of the amygdala, the prefrontal cortex (involved in reasoning and emotional control), the VTA, and the NA reduces stress-induced reinstatement.

CRAVING AND MIXED AFFECT

Although scientists have traditionally viewed craving as a distinct construct with important ties to affect, recent theorists have posited that craving is itself a form of affective response that reflects an approach-oriented state, similar to how fear is an avoidance-oriented state. Yet, addicts report that craving is a predominantly aversive experience, and research reveals that it is typically correlated with self-reported negative affect. Although craving appears to be universally associated with heightened arousal, context appears to moderate the valence of the craving experience. The availability of the drug may play an important role in determining the valence of craving, as research with cigarette smokers indicates that negative affect is substantially lower following the induction of craving if individuals believe they can smoke immediately afterward. One plausible explanation is that when drugs are immediately available the addict may

anticipate relief, which may be experienced as an appetitive state. In contrast, when drugs are not available, craving may be a frustrating experience and take a more negative tone.

Theories of Craving and Affect

A number of prominent theories of drug addiction help explain how context can modify the affective valence of craving. According to Tiffany's Cognitive Model of Drug Urges and Drug-Use Behavior, drug use becomes automatic as dependence develops, and craving emerges as a result of these automated processes being disrupted. Therefore, while ongoing drug use in contexts where drugs are readily available would not have specific affective consequences in this model, the interruption of drug use will have distinct affective consequences. In situations where drug use is blocked by external circumstances (e.g. a supply of the drug running out), the addict would experience significant frustration and the overall affective response would have a distinctly negative tone. However, in situations where drugs are readily available but use is blocked by internal states (e.g. a desire to stop using), the affective consequences of craving would be better described as ambivalent, rather than strongly negative.

The Elaborated-Intrusion Theory of Desire purports that the specific affective response that craving may take will vary over the course of the experience. While an initial fleeting thought about a drug may be pleasant, this response can shift toward a distinctly negative response as this initial "intrusive" thought is more thoroughly processed and expanded upon. In some respects, this resembles the affective sequence characteristic of obsessive–compulsive disorder. Importantly, attempts to suppress thoughts related to drug use, particularly in situations where the drug is unavailable (e.g. the individual cannot afford to purchase the drug) or drug use is undesirable (e.g. the person is trying to stop using), will paradoxically lead to further cognitive elaboration of drug-related thoughts. At this stage, craving begins to take a predominantly negative tone. Although this negative tone may persist for some time, craving may return to a positive state at times as elaboration of the drug-related thought incorporates memories of positive sensory experiences resulting from past drug use. These intermittent bouts of positive states may serve to maintain the state of craving, even in instances where craving is predominantly negative. Deprivation and withdrawal may also play a key role in this process, by increasing the likelihood of intrusive thoughts and enhancing the intensity of the negative affect experienced. Also critical to this model, the relationship between intrusive thoughts and deprivation is thought to be reciprocal, as intrusions may serve to highlight awareness of deprivation and further increase the motivation to use drugs to mitigate any consequences resulting from that deprived state.

CONCLUSIONS

A great deal of research has demonstrated the critical role that withdrawal, craving, and affect play in the onset, development, and maintenance of substance use and addiction. While defining these constructs has proven exceptionally challenging, it is clear that these constructs intersect in a number of different ways. It is this intersection that forms the core of many prominent models of addiction. This chapter has described several ways in which these concepts interact, including the impact of substance deprivation on craving and affective experience. Even in the absence of acute substance deprivation, sensitization of neural reward pathways plays an important role in the generation of craving and drug use behavior, by heightening the reward value of drugs relative to other rewards. Certain individuals may be more prone to these facets of substance use, due to differences in affective and cognitive processes. Drugs may take advantage of these individual differences, speeding the development of dependence and increasing its severity. However, even in the absence of direct drug effects, the expected benefits of drug use may play an important role in the desire to consume substances of addiction. The motivation to use drugs is indeed a complex process, with affective correlates and consequences that can be moderated by a number of contextual factors.

The influence that research on deprivation, craving, and affect have had on shaping current views of addiction is profound and will likely continue to play an important role in the coming decades. Despite extensive research on the topic of addiction, it is clear that we have barely scratched the surface of a broad conceptual understanding of how addiction develops and is maintained. The majority of the relevant scientific literature has focused on examining individual components of the addictive process, often from the perspective of a single discipline. Recent advances in the fields of genetics, neuroscience, psychology, and others, coupled with an increased focus on transdisciplinary and translational research, will likely lead to important new discoveries, and an increased understanding of core components of addiction. More importantly, increased attention is now being given to how these components interact on neurobiological and behavioral levels to drive substance use and addictive behavior.

SEE ALSO

The Biopsychosocial Model of Addiction, Emotions and Addictive Processes, Self-Medication, Tolerance and Withdrawal, Implicit and Associative Processes in Addiction, Craving and Expectancies

Glossary

Brain-derived neurotrophic factor a protein promotes nerve cell growth and survival.

Corticotrophin releasing factor a hormone involved in the stress response.

Dispositional tolerance a decrease in the effects of a drug that result from a decrease in the amount of drug that reaches the site of action.

Functional tolerance a decrease in the effects of a drug that result from physiological changes that cause sites of action to be less responsive to a drug.

Negative reinforcement a process in which the occurrence of a behavior is followed by avoidance from an aversive stimulus, or the removal of a pre-existing aversive stimulus, resulting in an increase in the behavior.

Phasic craving an urge to use a drug that is evoked by internal or external stimuli – typically increasing and decreasing rapidly.

Positive reinforcement a process in which the occurrence of a behavior is followed by presentation of another stimulus, resulting in an increase in the behavior.

Tonic craving a state of craving that persists over long periods of time, tightly coupled with an individual's level of dependence and withdrawal status.

Further Reading

Baker, T.B., Piper, M.E., McCarthy, D.E., Majeskie, M.R., Fiore, M.C., 2004. Addiction motivation reformulated: an affective processing model of negative reinforcement. Psychological Review 111 (1), 33–51.

Evans, D.E., Drobes, D.J., 2009. Nicotine self-medication of cognitive-attentional processing. Addiction Biology 14, 32–42.

Gilbert, D.G., 1995. Smoking: Individual Differences, Psychopathology, and Emotion. Taylor & Francis, Washington, DC.

Goldman, M.S., 1999. Risk for substance abuse: memory as a common etiological pathway. Psychological Science 10, 196–198.

Kassel, J.D. (Ed.), 2010. Substance Abuse and Emotion. American Psychological Association, Washington, DC.

Kavanagh, D.J., Andrade, J., May, J., 2005. Imaginary relish and exquisite torture: the elaborated intrusion theory of desire. Psychological Review 112, 446–467.

Steele, C.M., Josephs, R.A., 1990. Alcohol myopia: its prized and dangerous effects. American Psychologist 45, 921–933.

Tiffany, S.T., 1990. A cognitive model of drug urges and drug-use behavior: role of automatic and nonautomatic processes. Psychological Review 97, 147–168.

Robinson, T.E., Berridge, K.C., 2003. Addiction. Annual Review of Psychology 24, 25–53.

Rottenberg, J., Johnson, S.L. (Eds.), 2007. Emotion and Psychopathology: Bridging Affective and Clinical Science. American Psychological Association, Washington, DC.

West, R., 2006. Theory of Addiction. Blackwell Publishing, Oxford, UK.

CHAPTER

42

Implicit and Associative Processes in Addiction

*Reinout W. Wiers** *and Alan W. Stacy*[$]

[*]University of Amsterdam, Amsterdam, The Netherlands [$]Claremont Graduate University, Claremont, CA, USA

OUTLINE

WHY DO THINGS THAT HARM?

Why do people engage in addictive behaviors such as smoking, binge drinking, or taking other addictive substances, while they "know" the consequences are harmful and potentially devastating? Research on implicit or associative processes in addiction has addressed this paradox and demonstrated that these choices are influenced by a set of associations in memory that become spontaneously activated, usually under specific conditions. These learned associations are not revealed through self-reflection, introspection, or reasoning processes, yet they guide behavior choices and the development of addictive habits. Hence, implicit cognition approaches have emphasized that there is more to decision making than a weighing of pros and cons of behavioral options. Implicit cognition may be especially powerful in the so-called "hot" situations, for example, when craving is high or in cases when control processes are impaired due to stress, fatigue, or acute alcohol intake. Note that although researchers on implicit cognitive processes in addiction emphasize the importance of these spontaneously activated associative processes, they also recognize that more rational decision making can influence drug decisions, especially when control capacity and motivation are sufficiently available.

ONE PROCESS OR MULTIPLE PROCESSES?

There is some discussion in the literature regarding how to best describe the cognitive processes involved in decision making, both in general and in relation to

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00042-5

addictive behaviors in particular. On one side of the spectrum there are theorists who state that there are two qualitatively different systems (or sets of systems), one reflective, rational, or propositional and one impulsive or associative, which independently influence (addictive) behaviors. On the other side of the spectrum are researchers who view expectancies, attitudes, or intentions as a unifying construct, capturing both associative and propositional processes. In the middle there are a variety of dual-process models, which distinguish between qualitatively different processes (associative versus propositional) and study the interaction of these processes. An influential general dual-process model was proposed by Strack and Deutsch, and many applications to addictive behaviors and proponents of these different perspectives can be found in the *Handbook on Implicit Cognition and Addiction* (see Further Reading). The study of associative processes in addiction can be undertaken both in animals, using controlled and well-studied learning paradigms, and in humans. Although we focus on human research here, we recognize the potential of the study of associative processes in addiction to bridge animal and human research.

DEFINITION

Definitions of implicit cognition vary depending on the discipline and research topic. Adapting a pragmatic approach that spans multiple disciplines (e.g. social and cognitive psychology), implicit cognitive processes are revealed on tests that do not require or encourage the conscious or deliberate recollection of previous events or introspections about the causes of one's behavior. Implicit cognition operates without reflective or deliberative processing or without being aware of the processes underlying behavior. There are many ways in which this nondeliberative process can be spontaneously triggered by a stimulus to influence behavior. Once triggered, an implicit process may influence the concepts that are activated in memory, one's subsequent train of thought, concomitant approach or avoidance tendencies, the accessibility of behavioral choices, and performance on tasks that reflect any of these consequences.

More specifically, processes may be considered implicit or automatic if they are characterized by one or more key qualities or features. De Houwer and colleagues suggest such characteristics include uncontrollability, absence of intentionality, goal independence, absence of awareness, efficiency, and operation even under time constraints. Implicit processes involve one or more of these qualities, suggesting a range of processes that are not necessarily unitary but nonetheless differ from reflective processes in important ways.

INDIRECT TESTS

Measures of implicit cognition use indirect assessment. Participants do not directly answer questions about the target construct. Indirect tests use measures of increased efficiency (e.g. reaction time or physiological measures), word production, or tests of memory performance that include an indirect element. The idea is that self-presentational or self-reflective processes are less likely to be engaged if direct questioning is not used. However, this is not absolute: research on faking of these indirect tests and research that mathematically models the underlying processes have indicated that both associative and controlled processes influence task outcomes. However, it is also clear that it is far easier to influence results on a direct than an indirect test, if a person is motivated to do so. Further, neurological approaches have shown strong evidence that indirect tests can detect associative cognition even when reflective processes are severely impaired and cannot be detected on direct tests.

DIFFERENT PROCESSES

Different implicit processes are distinguished. On the "input" (perceptual) side, an attentional bias for drug-related stimuli has been found for many different drugs, as well as for addictive behaviors not involving substances (e.g. gambling). On the "output" (motor) side, an action tendency to approach drug-related stimuli has been observed for a number of different substances (alcohol, marijuana, smoking, smaller research base as yet). Between input and output, memory associations are studied. As outlined in detail by Strack and Deutsch, these input, output, and mediating processes are viewed as intimately related from the dual-process perspective. Perception automatically triggers a motivational orientation and corresponding action tendencies, and the reverse also holds true: action tendencies can influence motivational orientation, which influence perception. Further, positive alcohol-related memory associations have been found to predict an attentional bias for alcohol, in participants with relatively modest control capacity. Regarding assessment and findings concerning a substance-related attentional bias, *see* Attentional Biases in Craving.

ASSESSMENT AND FINDINGS

Memory Associations

Memory associations can be assessed with a variety of measures that infer association strength on reaction

times in relation to different stimuli, and with a variety of memory association tests that do not rely on reaction times.

Measures of Associations Using Reaction Times

The Implicit Association Test (IAT), developed by Greenwald and colleagues, is the most commonly used RT-test to assess implicit associations. In addiction research, the test assesses individual differences in associations between a substance and two attribute categories (e.g. "positive" versus "negative"). The target category (alcohol or another substance in addiction research) also requires a contrast category (e.g. soft drinks or water for alcohol). On each trial of the task, participants rapidly categorize visually presented stimuli (pictures or words) by pressing one of two response keys. For example, in one phase of the task, they press the left response key when an alcohol-related or a positive stimulus is presented, and the right response key in response to alcohol-unrelated or negative stimuli. In another phase of the task, they press left when an alcohol-unrelated or positive stimulus is presented and the right key when an alcohol-related or negative stimulus is presented. The rationale for the task is that if participants automatically evaluate alcohol as positive rather than negative they should be quicker to respond when "alcohol" and "positive" stimuli share the same response key (as in the first phase in the example) compared to another block of the task where "alcohol" and "negative" words share the same response key (as in the second phase in the example). The IAT has a number of strengths: it is a flexible tool (different associations can be assessed), easy to use, and more reliable than many other implicit measures. Although the validity of the measure sometimes has been criticized, with ongoing debate, there are many studies revealing the test's utility, also in the field of addiction. New varieties of the IAT have been introduced, which may be particularly useful for the assessment of memory associations with substances for which it is hard to find a natural contrast category (e.g. smoking), because in these varieties there is no contrast category (Single Category IAT) or a contrast category that is not used in the instruction to categorize (Brief IAT).

Studies using varieties of the original IAT to assess whether alcohol is more strongly associated with negative than with positive effect found, perhaps surprisingly, consistently stronger associations between alcohol and negative valence than between alcohol and positive valence, in both light and heavy drinkers. However, heavy drinkers demonstrated somewhat less strong negative associations than light drinkers and IAT scores of implicit alcohol attitudes also predicted drinking behavior above the variance explained by explicit measures using the same words. Positive and negative associations can also be assessed separately, in two separate IATs contrasting positive with neutral and negative with neutral words. With these varieties, participants demonstrate both positive and negative associations for alcohol and smoking. Interestingly, the strength of the (relatively weak) positive associations and not the strength of the (relatively strong) negative associations predict unique variance in drinking and smoking behavior above the variance explained by explicit measures. Many studies have now found that relatively strong negative associations to be unrelated to drinking and smoking behavior. This suggests that positive associations may be more personally relevant, while negative associations may primarily reflect the common opinion of the culture or the fact that negative stimuli are generally more salient than positive stimuli. In line with this idea, studies using personalized versions of the IAT have demonstrated positive implicit associations with both alcohol and smoking. In addition to associations with positive and negative valence, researchers have used the IAT to assess other associations with substances, notably, associations with arousal, which were consistently found in heavy and problem drinkers, but not in light drinkers. Researchers have also recently used an IAT to assess automatically activated coping motives. The IAT can also be adapted to assess action tendencies to approach or avoid the substance (see below).

Other reaction time measures have been used to assess substance-related associations, such as varieties of priming measures. In these measures, a prime stimulus is presented briefly before the stimulus to which participants react (for example, by judging whether the second stimulus is a word or not). The advantage of this priming measure is that temporal order can be introduced to the measurement, which is not possible with the IAT and related measures. For example, it can be tested whether negative effect activates the concept of alcohol (or another substance) or whether the substance activates the concept of negative effect. With the varieties of these measures, it has been demonstrated that problem drinkers with high levels of psychiatric distress automatically activate the concept alcohol in reaction to negative effect primes.

Measures of Associations Not Using Reaction Times

Tests of memory associations using word production in addiction typically have used various types of word association tests. Common tests have used free word association, in which the participant lists the first word that comes to mind in response to a cue word, phrase, or picture, or a variant termed controlled

association, in which a category of some type (e.g. verb) is requested using similar instructions. Given that, tests do not directly inquire about the target concept (e.g. drug associations), these tests are indirect and may assess implicit processes. Indeed, consistent evidence across diverse paradigms from memory research shows that word association tests are capable of detecting implicit conceptual memory, and associations uncovered in these tests predict the spontaneous activation of cognitions across a wide range of experimental procedures. Probably the most compelling evidence for the implicit quality of word association tests comes from studies in amnesic participants, who show severely impaired explicit memory but no impairments on tests of implicit memory.

Addiction research began using word association methods comprehensively with the work of Szalay and colleagues. They found different associative structures in drug users versus nonusers and among participants entering versus successfully completing drug treatment. Most of the addiction research using word association conducted after this pioneering work has relied on traditional word association, using either free word association or a form of controlled association termed verb generation; verb generation asks for the first action or behavior that comes to mind in response to the cue, which may be a word, picture, or other stimuli. In the first study using this technique, Stacy and colleagues asked college students to generate the first behavior that came to mind in response to a series of alcohol-related and neutral short phrases. The alcohol-related phrases did not explicitly mention alcohol or its synonyms, but were obtained from college student norms for likely (perceived) positive outcomes of alcohol use (e.g. having fun, feeling good). Strong correlations were found between the generation of alcohol responses and alcohol consumption, even though nothing was asked about alcohol until after the associations were elicited using this indirect assessment. A number of studies have replicated this finding but have also documented the importance of associations between cues (in addition to affective associations) and alcohol as well as other substances. In a meta-analysis by Rooke and colleagues, word association tasks demonstrated the best predictive effects among all indirect tests of alcohol or other drug-related associations studied to date.

Action Tendencies

Motivational theories distinguish between two motivational orientations in relation to specific stimuli: approach and avoidance. Note that a wider perspective on motivation, proposed by Cox and Klinger, also involves broader concerns relevant to substance use, such as other personal concerns (e.g. partner relationships, work, housing, etc.), which may also be assessed with indirect measures (but little work has been done yet to achieve this). We focus here on the more narrow motivational orientation (the tendency to approach alcohol/drug cues). The first is an adapted IAT, where associations with approach versus avoidance are measured by comparing reaction times in a phase where the substance is categorized together with approach-words, with reaction times in a phase where the substance is categorized together with avoid-words. The second measure uses a symbolic approach or avoid movement by a manikin (matchstick Figure), which has to be moved toward the substance in one phase and away from the substance in another phase (reaction times are compared again). The third task involves an actual approach or avoid movement with a joystick, and participants respond to a feature of the stimulus that is unrelated to the contents (e.g. pull joystick in response to pictures tilted to the left, push in response to pictures tilted to the right). With the first measure it has been demonstrated in several studies that heavy drinkers are faster to categorize alcohol words when the same response key is used to categorize approach words than when the same response key is used to categorize avoid words; this facilitative effect on alcohol words is predictive of participants' self-reported craving when faced with an alcoholic beverage. With the symbolic approach avoid task (sometimes referred to with the too general label SRC, which stands for stimulus response compatibility), it has been found that frequent substance users (of alcohol, tobacco, or marijuana) are faster to make the manikin approach of the substance pictures than to make the manikin avoid the substance pictures. This approach bias has been related to indices of the severity of their substance use. In the third assessment task, the approach avoidance task (AAT), participants respond to pictures (substance or not substance-related) presented on a computer screen, by pulling or pushing a joystick. The task incorporates a "zooming mechanism": when the joystick is pulled, the picture size increases on the computer screen, and when it is pushed, it decreases. This zooming mechanism creates a sensation of approach or avoid and disambiguates the task. In several studies the AAT has been used in a so-called irrelevant feature variety, which implies that participants react to a feature of the picture that is unrelated to the contents of the picture (e.g. picture format: landscape or portrait, or pictures are presented with a small tilt to the left or right). For example, participants are instructed to respond with an approach movement (pull the joystick) to all pictures in portrait format and with an avoid movement (push the joystick) to all pictures in landscape format. The pictures in the task include control pictures (e.g. soft drinks, and sometimes also other pictures like landscapes) and critical stimuli (e.g. picture of beer bottle);

the type of picture is not addressed in the instructions. With this variety of the AAT it has been found that heavy drinkers as well as alcoholic patients are faster to approach than to avoid alcohol stimuli. An approach (or avoidance) bias is generally not found for control pictures that are positive (or negative). This could be related to the natural coupling of the stimulus category (appetitive stimuli) to the response used (arm movement). Note that the AAT is more indirect compared to the two tasks discussed previously (IAT, SRC), which both involve the instruction to categorize the substance pictures with approach in one phase of the test and with avoid in another phase. Initial findings of an approach bias for other substances (cigarettes and marijuana) have been reported as well, with an approach bias for marijuana as the best predictor of escalation six months later in problematic users. In addition, training varieties of the AAT have been used recently to change the approach bias for alcohol (see below).

Individual Differences: Trait and State

As noted in the introduction, dual-process models state that many behaviors, including addictive behaviors, are influenced both by automatically activated associative processes and by more controlled reflective processes. The relative extent to which associative and reflective processes determine (addictive) behaviors is likely to differ between people (trait differences) and within a person, depending on the situation (state).

Trait

In a series of studies using different measures of substance associations (open-ended memory associations as well as reaction time measures) and different measures of individual differences in executive control, it was demonstrated that associative measures are a better predictor of substance use in individuals with relatively weak scores in executive control, than in individuals with good scores in executive control. There is also some (more limited) evidence that in individuals with relatively good scores in executive control, explicit expectancies better predict substance use. This dissociation has not only been found repeatedly for substance use outcome measures, it has also been found for related behaviors, such as aggressive tendencies, aggression after alcohol, sexual interest, and candy eating despite attempts to restrain eating. In all of these cases, the target behavior was better predicted by measures of relevant associations in individuals with relatively weaker scores in executive control.

State

There is also an emerging evidence that measures of the relevant associative processes better predict behavior in conditions where the (limited) executive control capacity has been reduced. This can be done by depleting executive control. One way to do this is to first give participants a task that depletes control capacity, such as holding a hand in ice water or by doing a number of difficult cognitive tasks. There is also evidence that an acute dose of alcohol temporarily impairs executive control capacity. As reviewed by Field and colleagues, alcohol not only impairs executive control but it also primes associative appetitive processes. Depending on the situation, this can lead to the consumption of more alcohol than desired before the drinking session, or to an increased chance of performing other behaviors than normally desired, including aggression, overeating, and unsafe sex. As reviewed by Wiers, Houben, and colleagues, this may explain why in some cases after alcohol consumption, common sense may go out of the window. As yet, there is little research on the effects of other substances on associative processes, on the one hand, and executive control processes, on the other hand. In addition, there is little research yet on long-term effects of different substances on these processes, as well as on other relevant cognitive motivational processes (e.g. motivation to engage in school or work). The suggestion from clinical observations as well as animal research is that negative long-term effects are likely.

ORIGINS AND RELATION TO THE DEVELOPMENT OF ADDICTIVE BEHAVIORS

As yet little is known about the origins of implicit cognitive processes involved in addictive behaviors in humans. As reviewed by De Wit and Dickinson, in animal research, there is a rich tradition on different types of associative learning (classical and instrumental conditioning) with recent emphasis on different underlying neural systems involved. Some of these tasks have recently been used in humans as well. Most human research on the origins of associative processes in addiction have used natural existing drug-related stimuli, rather than studying reactions to novel artificial stimuli of which the learning history is controlled, as is done in the animal research. In these studies of natural origins of alcohol- and drug-related association, one important distinction concerns origins before personal initiation of the substance and origins after initiation.

Origins Before Substance Initiation

Some of the most likely origins of implicit associations involve the social environment (parents, peers) and media exposure. Grenard and colleagues found

that the growth in alcohol-related implicit associations in adolescence over time is predicted by earlier reactions to alcohol advertising on television. Alcohol consumption and implicit association growth curves were strongly correlated across early adolescence. There is also some evidence indicating that parents' implicit associations with smoking are related to children's implicit smoking associations, as well as with their smoking behavior. Traits like sensation seeking and impulsivity have also been related to the early initiation of substance use, but to our knowledge it has not been studied to what extent this relationship is mediated by substance-related cognitions.

Origins After Substance Initiation

It is a well-established finding that addictive behaviors are partly explained by genetic factors (*see* Neurogenesis and Addictive Disorders on Neurogenetics of Addiction). There are some first indications that some of these genetic factors may be mediated by cognitive processes. For example, it has been demonstrated that heavy drinkers with a g-allele in the *OPRM1* gene (also associated with strong cue-induced craving) have stronger automatic approach tendencies to alcohol cues than equally heavy drinkers without this allele. Other genetic variations are currently studied in relation to associative processes in addiction, as well as in relation to other psychopathology (e.g. an attentional bias for emotional stimuli in anxiety has been related to varieties of the serotonin transporter gene promoter region (5-HTTLPR) Polymorphism).

In relation to neurobiological models of the development of substance use, some implicit processes (attentional bias, arousal associations, and approach tendencies) have been theoretically linked to the development of a hypersensitive motivational reaction to conditioned stimuli predicting drug use (incentive sensitization, *see* Incentive Salience and the Transition to Addiction). In addition, it has been proposed that some associative processes may be more related to the development of drug-related habits (e.g. cue–behavior associations). Both hypotheses await critical testing.

Development and Causal Role

As yet little is known about the development of implicit cognitive processes in addictive behaviors.

Regarding a causal role of implicit cognitive processes, the current database is limited. It is well established that substance associations differ between heavy and light users or abstainers of the substance and that implicit measures predict variance in the addictive behavior after controlling for explicit predictors (as demonstrated in the meta-analysis of Rooke and colleagues). There is also some evidence of prospective short-term prediction after controlling for previous substance use and explicit cognitions, but this database is more limited. A more direct way to test the causal status of implicit cognitive processes is to directly manipulate them, and to study the effects on the cognitive processes and on behavior. Since these recently developed methods also have therapeutic implications, they are discussed below.

POTENTIAL FOR NEW INTERVENTIONS

Addiction researchers have begun attempts to directly manipulate implicit cognitive processes. This can be done for two reasons: to establish the causal status of the process in an experimental way and for therapeutic reasons. In studies establishing the causal status, it is as informative to experimentally increase a cognitive bias as it is to decrease the same bias and to observe corresponding effects on addictive behaviors. For ethical reasons, these studies are typically performed in nonaddicted participants. In clinical samples, the cognitive substance bias is typically decreased in the experimental group and left unaffected in the control group. The manipulation or retraining of the bias is typically done by using a variety of the task used to assess the same bias, in which a contingency is built in. For example, in an assessment version of the AAT, participants respond to an equal number of alcohol pictures with an approach movement and with an avoidance movement (the difference generates the approach bias for alcohol). In a manipulation version, this contingency is changed; in an avoid-alcohol condition, participants are made to respond with an avoidance movement to most alcohol pictures and with an approach movement to most nonalcohol pictures. Similarly, in an approach-alcohol condition, participants are made to respond with an approach movement to most alcohol pictures, and with an avoid movement to most nonalcohol pictures (not used in clinical studies). In a control condition, often continued assessment is used (equal number of approach and avoid movements to alcohol and nonalcohol pictures) or no assessment. Researchers study the effects of this manipulation on pictures used during the training, as well as to pictures not used during the training (close generalization; same task as used in the training, but different stimuli). They also study further generalization (e.g. performance on another task that attempts to assess the same process), and effects on behavior.

Following the pioneering work in anxiety research, addiction researchers first targeted an attentional bias for a substance (alcohol). Initial studies testing causality by a single session of attentional retraining found corresponding effects on the bias, but no generalization to untrained stimuli or different tasks. However, recent

studies using repeated attentional retraining in problem drinkers and alcoholic patients found generalized effects on the attentional bias and reduced alcohol intake in problem drinkers after training in one study and increased levels of abstinence in alcoholic patients in another better controlled study by Schoenmakers and colleagues. Wiers, Rinck, and colleagues developed a retraining version of the alcohol AAT. In a preclinical study, they found generalized effects on the approach bias for alcohol in students (stronger approach bias in students who had responded with an approach movement to most alcohol stimuli, and reduced approach bias in students who had responded with an avoidance movement to most alcohol stimuli). In students who were successfully trained, they also found corresponding behavioral effects in a taste test: students who had been successfully trained to make a push response to alcohol stimuli, drank less on a subsequent taste test involving different beers. In the first randomized controlled clinical study using this procedure, alcoholic patients were trained in four sessions to avoid alcohol and approach nonalcoholic drinks (experimental condition) or not (control condition), after which the patients participated in regular cognitive behavioral therapy (CBT, *see* Cognitive Behavioral Therapies). This resulted in a strong generalized effect on the automatically activated tendency to approach alcohol in the experimental condition. For example, patients who had been trained to make a push-movement with the joystick in response to alcohol pictures, also associated alcohol words more strongly with avoidance than with approach on an IAT, while patients in the control group continued to associate alcohol words more strongly with approach-words than with avoidance-words. Moreover, patients in the experimental condition were 13% more likely to be abstinent 1 year later than patients in the control condition.

In addition to first experimental studies targeting an attentional bias and approach bias, researchers have initiated attempts to change substance-related associations. This can be done by consistently coupling the substance to negative (repulsive) pictures, a form of evaluative conditioning. In two studies, Houben and colleagues coupled alcohol pictures with negative pictures (experimental condition), which was not done in the control condition (alcohol pictures were coupled to neutral pictures). In both studies, participants in the experimental condition drank less in the week after the experiment than before, which was not the case in control participants. In one of the studies, it was also shown that the experimental condition had stronger negative alcohol associations after the experiment than the control condition (using an IAT). To the best of our knowledge, evaluative conditioning has not yet been used in clinical samples, but it bears resemblance to aversive conditioning procedures used earlier in the treatment of alcoholism (unfortunately often in poorly designed research).

Note that it is well possible that existing CBT may result in changes in substance associations, but these are often not measured in studies of CBT effects. One study that did assess the effects of a specific form of CBT (a so-called expectancy challenge) on alcohol associations (in addition to explicit expectancies) found that while this intervention resulted in the anticipated changes in explicit expectancies, there was little change in alcohol associations. However, there may also be other ways to use explicit instruction to change the likelihood of the desirable over the undesirable behavior occurring in a specific situation. This may be done, for example, by forming the so-called implementation intentions: if situation X, then behavior Y. For example, if I drive, then I order nonalcoholic drinks.

CONCLUSION

There is ample evidence that implicit or associative processes play an important role in addictive behaviors, despite the fact that there are many open questions in this relatively young field of research. There are different indirect methods to assess these processes, some relying on reaction times, some not. With both types of measures, it has been found that substance associations predict unique variance in substance use, also after controlling for explicit cognitive processes. In addition, there is evidence that heavy substance use is related to an approach bias for that substance (as well as an attentional bias for the substance, discussed more extensively in Attentional Biases in Craving). Most research has studied associative processes involved in alcohol use and misuse, some in cigarette and marijuana smoking and few have studied associative processes related to other substance use and misuse. In addition, little is known yet about the development of associative processes in relation to the development of addictive behaviors. There is an emerging evidence that associative processes better predict substance use and related behaviors in individuals with relatively limited executive control abilities (either as a trait or as a state). Finally, there are promising first studies that demonstrate that it is possible to modify associative processes in addiction. These findings are not only important for establishing causality of these processes but also hold promise for novel interventions, as initial studies demonstrate.

List of Abbreviations

AAT	approach avoidance task
CBT	cognitive behavioral therapy
5-HTTLPR	serotonin transporter gene promoter region
SRC	stimulus response compatibility

Glossary

IAT Implicit Association Test, a reaction time test that aims to assess the associative strength between concepts.

Further Reading

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., Moors, A., 2009. Implicit measures: a normative analysis and review. Psychological Bulletin 135 (3), 347–368.

de Wit, S., Dickinson, A., 2009. Associative theories of goal-directed behaviour: a case for animal–human translational models. Psychological Research 73 (4), 463–476.

Field, M., Wiers, R.W., Christiansen, P., Fillmore, M.T., Verster, J.C., 2010. Acute alcohol effects on executive function and implicit cognition: implications for loss of control. Alcoholism Clinical and Experimental Research 34 (8), 1346–1352.

Houben, K., Havermans, R.C., Wiers, R.W., 2010. Learning to dislike alcohol: conditioning negative implicit attitudes toward alcohol and its effect on drinking behavior. Psychopharmacology (Berl) 211 (1), 79–86.

Roefs, A., Huijding, J., Smulders, F.T.Y., MacLeod, C.M., De Jong, P., Wiers, R.W., Jansen, A.T.M., 2011. Implicit measures of association in psychopathology research. Psychological Bulletin 137, 149–193.

Rooke, S.E., Hine, D.W., Thorsteinsson, E.B., 2008. Implicit cognition and substance use: a meta-analysis. Addictive Behaviors 33 (10), 1314–1328.

Schoenmakers, T., Lux, I., Goertz, A., Van Kerkhof, D., De Bruin, M., Wiers, R.W., 2010. A randomized clinical trial to measure effects of an intervention to modify attentional bias in alcohol dependent patients. Drug and Alcohol Dependence 109, 30–36.

Stacy, A.W., Wiers, R.W., 2010. Implicit cognition and addiction: a tool for explaining paradoxical behavior. Annual Review of Clinical Psychology 6, 551–575.

Stacy, A.W., Ames, S.L., Wiers, R.W., Krank, M.D., 2010. Associative memory in appetitive behavior: framework and relevance to epidemiology and prevention. In: Scheier, L.M. (Ed.), Handbook of Drug Use Etiology: Theory, Methods, and Empirical Findings. APA Books, Washington, DC.

Strack, F., Deutsch, R., 2004. Reflective and impulsive determinants of social behavior. Personality and Social Psychology Review 8, 220–247.

Szalay, L.B., Bovasso, G., Vilov, S.K., Williams, R.E., 1992. Assessing treatment effects through changes in perceptions and cognitive organization. American Journal of Drug & Alcohol Abuse 18, 407–428.

Wiers, R.W., Bartholow, B.D., van den Wildenberg, E., et al., 2007. Automatic and controlled processes and the development of addictive behaviors in adolescents: a review and a model. Pharmacology, Biochemistry and Behavior 86, 263–283.

Wiers, R.W., Eberl, C., Rinck, M., Becker, E., Lindenmeyer, J., 2011. Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. Psychological Science 22 (4), 490–497.

Wiers, R.W., Houben, K., Roefs, A., De Jong, P., Hofmann, W., Stacy, A.W., 2010. Implicit cognition in health psychology: why common sense goes out of the window. In: Gawronski, B., Payne, K. (Eds.), Handbook of Implicit Social Cognition. Guilford, New York, pp. 463–488.

Wiers, R.W., Stacy, A.W. (Eds.), 2006. Handbook of Implicit Cognition and Addiction. SAGE Publishers, Thousand Oaks, CA.

Relevant Websites

http://www.implicit.harvard.edu/implicit/– Educational resource and research site for investigations in implicit social cognition

http://www.adaptlab.eu– Adapt – Addiction, Development and Psychopathology (ADAPT) lab, Department of Psychology, University of Amsterdam

CHAPTER

43

Cue Reactivity

*Abigail K. Rose**, *Matt Field**, *Ingmar H.A. Franken*[$] and *Marcus Munafò*[**]

[*]University of Liverpool, Liverpool, UK [$]Institute of Psychology, Erasmus University Rotterdam, Rotterdam, The Netherlands [**]University of Bristol, Bristol, UK

OUTLINE

CUE REACTIVITY

Rates of relapse across all classes of addictive substance are high (substance covers alcohol and nicotine as well as illicit substances; see Fig. 43.1). For example, up to 70% of alcohol-dependent individuals relapse within 3–6 months of receiving treatment, while up to 90% of attempts to quit smoking fail within a year. Relapse rates for opiates, 12 months following residential treatment, are approximately 65–70%. A risk factor believed to be important in substance-related behaviors (defined here as substance-seeking and self-administration) is cue reactivity. When an individual comes into contact with a stimulus (cue, e.g. alcohol advertisements, the sight and smell of cigarettes, other substance users) associated with the substance, psychological and physiological reactions can occur, which may trigger substance-related behaviors. Substance cues can increase self-reported craving and self-administration behaviors in both dependent and nondependent populations and the degree of cue reactivity may predict individual differences in the risk of relapse. However, the evidence is inconsistent: cue reactivity is not always observed, and the predicted relationships between cue reactivity and other variables, such as severity of

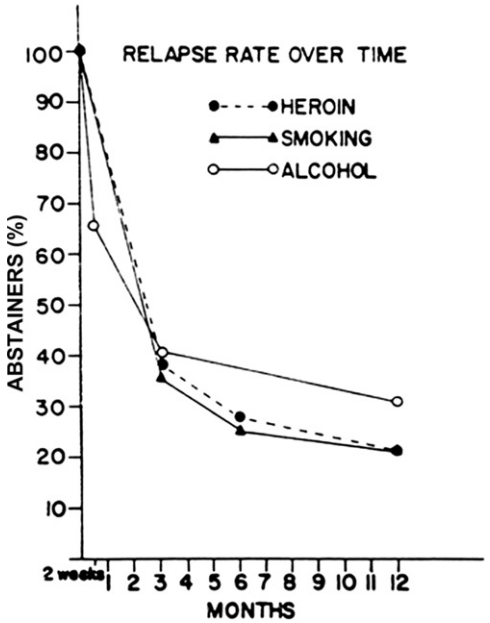


FIGURE 43.1 Relapse rate over time for heroin, smoking, and alcohol. *Reproduced with permission from Hunt, W.A., Barnett, L.W., Branch, L.G. 1971. Relapse rates in addiction programs. Journal of Clinical Psychology, 27, 455–456.*

http://dx.doi.org/10.1016/B978-0-12-398336-7.00043-7

dependence or prospective substance use, are not always present.

Cue reactivity paradigms measure the effect of both discrete (e.g. a glass of wine or a tone paired with substance administration) and contextual (e.g. the pub, operant chamber) cues on a range of measures. Cue reactivity effects have been found on measures of self-report (e.g. desire and craving), automatic cognitive responses (e.g. approach tendencies, attentional bias), self-administration (e.g. ad lib administration, speed of consumption), motivation (e.g. operant responding), physiological (e.g. heart rate, skin conductance) and pharmacological (e.g. dopamine) responses, and increased activation in discrete regions of the brain (e.g. ventral striatum, anterior cingulate cortex).

As substance dependence is necessarily preceded by a period of time in which the substance is administered on a regular basis, often in predictable environments and when in the presence of discrete cues, conditioning-based accounts of cue reactivity have been highly influential in this area. However, other cognitive theories have provided different insights and identified several cognitive biases that are likely to be involved in cue reactivity. As the neurobiological basis of cue reactivity has been mapped, it is clear that reactivity involves a complex system of interrelated processes. Finally, it is important to gauge how cue reactivity can be applied to clinical aspects of addiction and inform treatment.

CONDITIONING AND LEARNING THEORIES

It was originally suggested that, as dependence on a substance developed, withdrawal symptoms (unconditioned responses) would be experienced and cues (conditioned stimuli) associated with substance administration would come to elicit withdrawal-like responses (conditioned responses; see Fig. 43.2). Substance use is, in this theoretical account, defined as an operant behavior that alleviates withdrawal symptoms (avoidance or escape learning), and each successful reduction of negative effect reinforces substance use. These withdrawal-based, negative reinforcement theories were originally formulated due to anecdotal reports from opiate addicts who claimed to experience withdrawal-like symptoms when coming into contact with opiate-related cues (e.g. previous substance-related contexts).

Both human and animal models have shown that, if withdrawal is accompanied by a conditioned stimulus (e.g. a tone or odor), the conditioned stimulus alone can precipitate withdrawal. However, cues associated with substances have also been found to alleviate withdrawal symptoms. Although the ability of cues to trigger withdrawal symptomatology is important, the key issue is whether this is related to maintenance of problematic substance use and relapse. There is a substantial amount of research that fails to find these relationships. For example, some addicts never abstain long enough for conditioned withdrawal to develop yet they persist in self-administering substances. In addition, anecdotal reports suggest that only a minority of heroin addicts experience conditioned withdrawal, and even fewer report relapse as a result of this effect.

In a similar vein, it has been proposed that, through attempts to maintain homeostasis, opponent processes (physiological and affective responses that work in direct opposition to the effects of the substance itself) develop in anticipation of, and to counteract, the effects of the substance (see Fig. 43.2). These opponent processes may underlie the development of tolerance and support the administration of greater substance doses to experience the desired effects. If opponent processes can be conditioned, substance cues associated with the substance's central effects could trigger the opponent process and reduce the perceived effects of the substance, or be perceived as withdrawal in the absence of substance administration. The absence of a conditioned opponent process has been put forward as a reason for why fatal overdoses occur in experienced substance users when they have administered a substance in an environment free from the usual substance cues. Importantly, both the theories of conditioned withdrawal and conditioned opponent processes argue that physiological changes are key in activating substance urges, and that these urges motivate substance intake to reduce these physiological changes. However, as with conditioned withdrawal, evidence for opponent processes is also lacking and research fails to show a strong relationship between physiological changes in response to cues and self-reported craving for substances.

Competing theories postulate that cues take on positive incentive properties and trigger substance-like effects (see Fig. 43.2). General incentive motivational frameworks propose that cues can develop conditioned incentive properties in their own right and elicit motivational states. These motivational states may support specific types of behavior and can interact with internal states. For example, the sight of a cold can of beer may elicit a desire to drink alcohol, which triggers approach and consummatory behaviors; however, this effect may be greater in a person who is also thirsty. In terms of withdrawal, instead of negative reinforcement *per se*, the withdrawal state makes the incentive value of the substance so great that substance use prevails. When applying these general ideas to addiction, substance-associated cues could elicit substance-like, as opposed

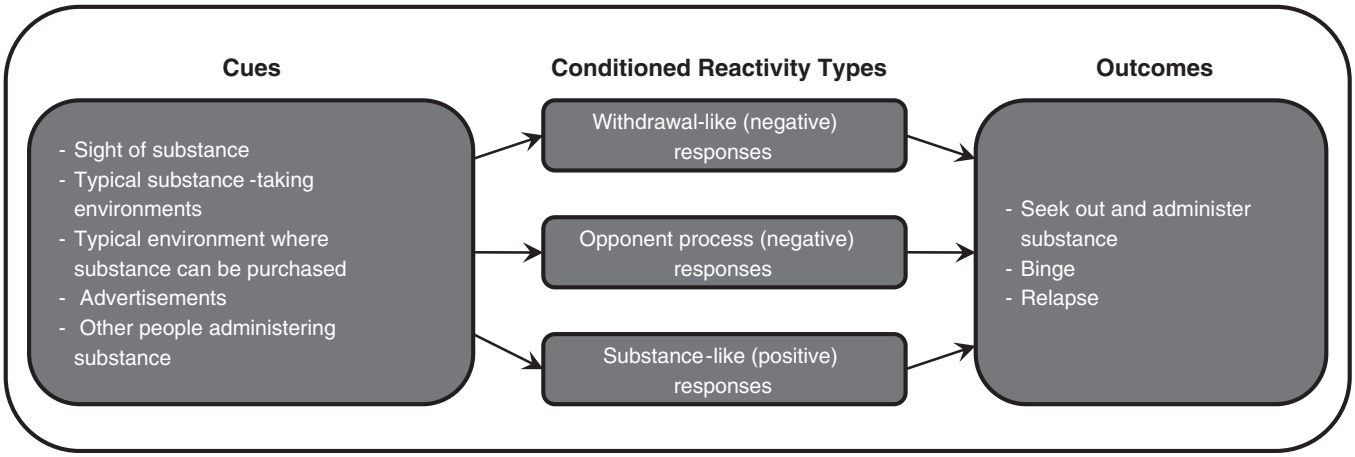


FIGURE 43.2 Typical types of cues, possible cue reactions according to major conditioning theories of addiction, and potential outcomes from cue exposure.

to substance-opposite, effects. Animal models have found that cues associated with opiate administration can produce hyperthermia, which mimics the actual substance effect, rather than hypothermia, which is a withdrawal effect.

The evidence concerning single process theories (i.e. those that focus on positive or negative reinforcement, and substance-like or substance-opposite effects) with regard to cue reactivity shows inconsistencies between direction of cue effect and differences in effect sizes across substance classes. For example, significant decreases in skin temperature reactivity have been found in opiate and cocaine addicts but not in alcoholics and dependent smokers. Self-reported desire in the presence of substance cues often increases significantly across all substance types, but effect sizes are inconsistent for opiate- and smoking-dependent populations.

Another key area addressed by learning theories is whether substance behavior is habit-like or goal-directed. Habit-like behavior is thought to be based upon stimulus–response associations, in which behavior (e.g. substance intake) is triggered by a cue with little or no mediation by the intention to engage in substance use, or anticipated outcomes of substance use. Goal-directed behavior involves stimulus–outcome–response associations, in which the cue triggers an expectancy of the outcome, which then triggers behavior. Devaluation paradigms are often used to compare these two learning processes and some animal models have indicated that substance rewards act differently to natural rewards. When food or substance outcomes have been devalued (e.g. through being paired with an aversive consequence or state-specific satiety), some research has found that animals will stop responding for the former but not the latter. These findings indicate that, whereas food-seeking behavior is goal-directed (i.e. governed by an expectancy of the outcome), substance-seeking behavior is insensitive to the devaluation effect, indicating a habit-like stimulus–response association. However, human research has yielded somewhat different results. Nicotine devaluation, by satiety, reduces cigarette-responding, as would be expected by a goal-directed theory, but the presence of a cigarette cue abolishes this devaluation effect and substance-seeking responses occur regardless of the substance's incentive value. This research demonstrates the Pavlovian-to-instrumental-transfer (PIT) effect in cue reactivity; conditioned stimuli (traditionally associated with stimulus–reward associations) for a given reward can elicit operant responding for that reward (response–outcome associations). When a substance cue is presented in extinction (no response outcome), the behavioral response (substance-seeking) is elicited, suggesting that the cue activates an expectancy of a specific outcome (e.g. alcohol) but not necessarily the current incentive value of that outcome.

Such findings indicate that although substance-related behavior involves both goal-directed and habit-like learning, it may also be particularly susceptible to the influence of cues. This possibility lends itself to the compulsive nature of addiction and fits with findings that cues can trigger relapse in individuals, even though they are aware of the significant adverse consequences of returning to substance use. Such possibilities may predict that treatments that emphasize the negative consequences of substance use may be limited in their efficacy. In addition, the finding that cues can trigger substance-related behavior, with scant regard to the value of the substance, may help explain the claims that craving elicited by cues is not consistently related to relapse risk. The effects of cues on substance behaviors may bypass craving mechanisms, which may be more associated with the perceived value of the substance, and trigger administration behaviors in a more automatic fashion. However, systematic

investigation of the relationship between cue-induced craving and relapse is still needed to resolve this issue.

Due to the discrepancies in results, theoretical accounts of substance use moved toward dual process theories with an increased focus on more complex learning processes and, initially, there were two primary versions: motivation and expectation. The motivation framework suggests that cues (e.g. an exteroceptive stimulus or an interoceptive state) can motivate a response and that the response outcome (e.g. a new state) can provide feedback to strengthen this association. Alternatively, the expectation framework argues that the cue first activates an expectation of the response outcome, which then triggers the response. The actual response outcome can then feedback on to the expectation (see Fig. 43.3). As the latter model incorporates outcome expectancies, the expectation framework better fits with the finding that if the outcome is devalued or revalued, the response for that outcome is decreased or increased, respectively. However, the model does not directly explain the possibility that cues may trigger a general outcome expectancy, which does not take into account the current value of the outcome.

Although most recent major theories of substance dependence acknowledge a role of conditioning, not all theories assume that conditioning is sufficient to explain substance use and relapse. In its simplest form, conditioning theories argue that over time, cues can elicit physiological responses and/or motivational states (e.g. urges, cravings), which promote substance use. However, there is a lack of consistent evidence that self-reported urges or physiological reactivity account for a significant amount of the variance seen within actual substance use. Severity of dependence is not always correlated with degree of cue reactivity, as would be predicted by a conditioning account, and not all dependent individuals experience cue reactivity. In addition, substance use, whether as an example of "everyday usage" or relapse, involves a number of aspects. For example, deciding to purchase a bottle of wine while shopping may take into account a number of factors including price, preexisting plans, substance-related memories, and, in the individual trying to abstain, perceived self-control.

COGNITIVE THEORIES

At the more basic level, it has been suggested that urges simply reflect a state of physiological arousal and that the details of this state have been cognitively interpreted and labeled as "urges" by the individual. Any substance cue may trigger arousal and the individual may perceive the cue as associated with substance administration (e.g. a former heroin addict may take a walk in a park where he/she used to shoot up, this context may trigger both arousal and memories of past substance use). These two processes may result in the individual interpreting the arousal as craving, which triggers substance intake. These propositions come from a general "cognitive labelling theory," which argues that the cognitive and contextual/cue state that the individual is in is key to the interpretation of arousal. In addition, arousal can have effects on other processes, including focusing attention, and emotional influences on decision making, which could result in attentional bias for substance cues and disrupted rational decision making regarding substance use. Although no detailed arousal theory of addiction exists, it is fair to assume that the greater the physiological arousal, the greater

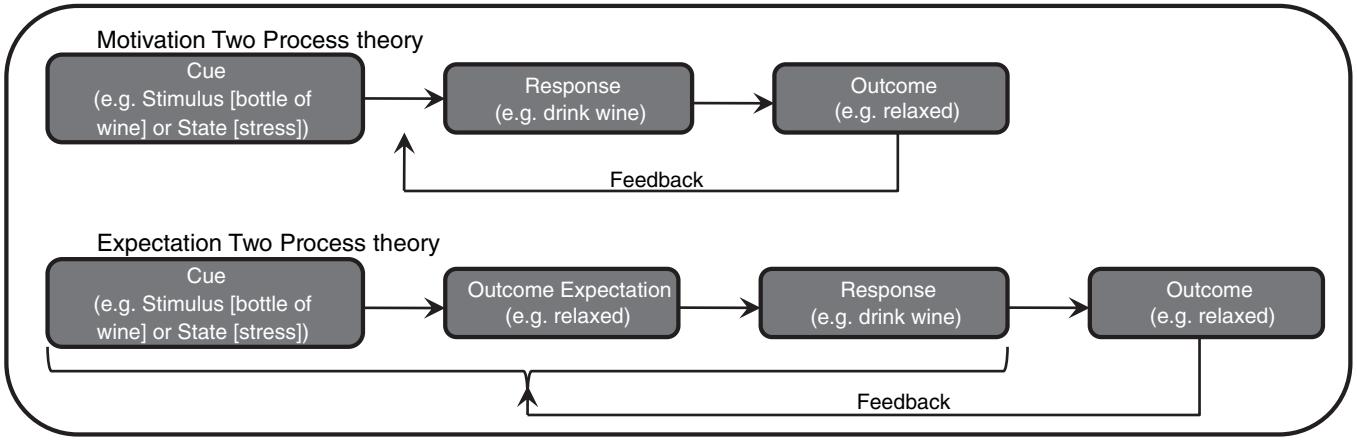


FIGURE 43.3 Outline of the motivation and expectation dual process theories. The motivation theory suggests that outcome of a response can feedback and strengthen the stimulus–response association. The expectation theory suggests that the cue stimulates an expectancy of the outcome, which then triggers a response (stimulus–outcome–response). The outcome therefore feedbacks on to this association and can affect the nature of responses made in future to the cue.

the perceived craving. However, as discussed above, clear relationships between physiological cue reactivity, craving, and substance intake have not been supported by the literature.

Some cognitive-behavioral theories argue that cues can trigger craving by highlighting the positive effects of the substance and, if powerful enough, this craving will translate into substance use. Outcome expectancies concerning the positive and negative reinforcing effects of substances are related to greater substance administration, more positive cue rating, and responding to cues with greater craving in dependent and non-dependent populations. An additional element to several cognitive theories is abstinence self-efficacy, which is the extent to which a person believes themselves capable of maintaining abstinence. Addicts with low abstinence self-efficacy are more likely to report higher levels of craving in the presence of cues. It is possible that the association between low abstinence self-efficacy and cue-induced craving reflects an impaired level of cognitive control, such that cues are able to have a greater effect on substance-related thoughts and behavior. Although this possibility needs further study, it is possible that low abstinence self-efficacy and cue reactivity may interact and promote relapse.

During the last two decades cognitive theories have become more complex. The "dual affect" theory argued that substance urges represent either positive affect, associated with appetitive motivational systems, or negative affect, related to withdrawal systems, and that these two frameworks of cognitive urges were mutually inhibitory. The dual affect theory proposed that substance administration, availability, and cues could activate a positive appetitive state and result in excessive and compulsive substance intake and polysubstance use. In addition, withdrawal, withdrawal-related cues, and aversive cues and states (e.g. stress) could trigger a "negative affect urge" framework and drive substance use. These urge frameworks were thought to be incorporated into larger propositional networks that contained information on salient cues and responses to these cues. The theory argued that as the conditions of the eliciting stimuli better matched the prototype, the stronger the network activation. In addition, as activation of some elements of the network occurs, the activation threshold for other propositional elements would be decreased. This suggestion can be related to cross-substance cue reactivity, for example, if an individual usually drinks, smokes, or takes cocaine while in a positive affective state, coming into contact with positively reinforcing alcohol cues will increase the likelihood of smoking and cocaine behaviors as well as drinking. There are a number of problems with the original dual affect theory. Evidence does not support mutually inhibitory affective systems, for instance, individuals who report positive reinforcing substance-related expectancies also tend to report high levels of negative reinforcing expectancies. In addition, increasing urge does not consistently increase measures of cue reactivity, and this lack of association is found both when measuring relationships within a substance class (e.g. urges for heroin may not increase heroin cue reactivity) and across substance classes (e.g. urges for alcohol do not always increase cigarette cue reactivity).

A prominent cognitive theory of addiction, the cognitive urge and automaticity model, proposes that, over time, behaviors associated with substance use become automatic (see Fig. 43.4). According to this theory, substance-related behaviors are stored in memory as "automatized action schemata" and these action plans contain information on behavioral sequences/procedures (e.g. how to open a bottle and then pour a glass

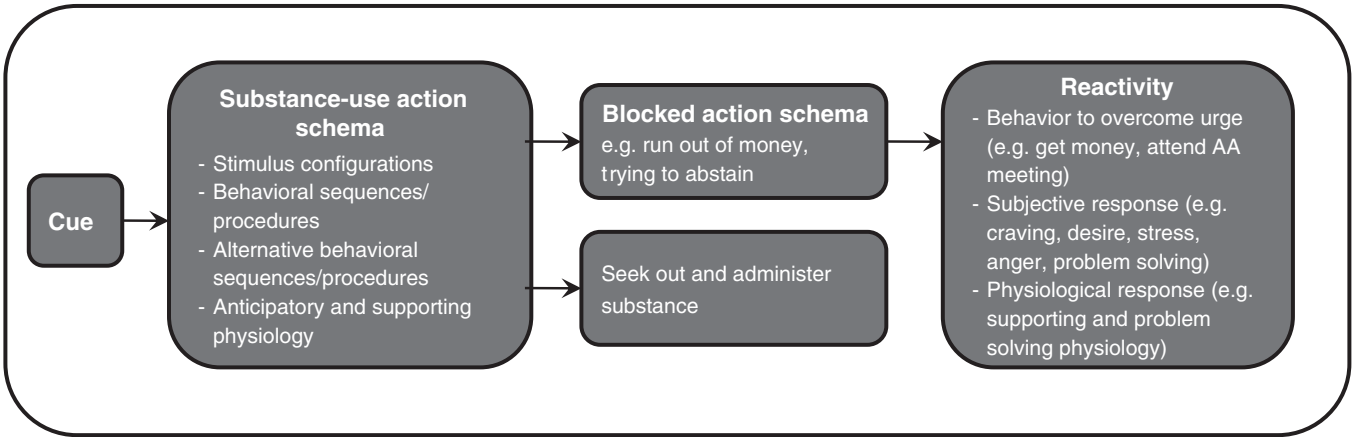


FIGURE 43.4 Cues can elicit automatic action schema which trigger substance-seeking and administration behaviors quickly and effortlessly. If these actions are blocked, effortful processes are activated to deal with this impediment and the individual may experience craving.

of wine, or find a vein before injecting heroin) and cues (e.g. a local pub, a state of stress, pharmacological effects of substance ingestion) which could trigger automatized behaviors (e.g. substance administration). In addition, the action plans incorporate physiological responses that support the behavior, and alternative behavioral sequences which could be triggered to ensure the behavioral goal is achieved even in the presence of obstacles (e.g. if a lighter fails to work, use matches). Repetition results in behavior becoming automatic, this is to say effortless, fast, and impervious to introspection. At a neuronal level, pathways are strengthened and activation thresholds decrease so that corresponding behaviors can be triggered easily and can occur quickly. Key to this model is that an urge reflects a nonautomatic process and only occurs if automatic substance-related behaviors are interrupted for some reason (e.g. no money, trying to abstain). These nonautomatic processes are the trigger for substance urges rather than the cue itself. This theory may be better able to account for studies which show that cue exposure fails to elicit reactivity as assessed by traditional measures (e.g. self-report, physiological reactions, behavior), although studies addressing this assumption are scarce. This cognitive theory is entirely consistent with other addiction models that emphasise habit-like processes, in the sense that cue-elicited effects are theorised to be relatively autonomous of the current motivational state of the individual.

BIASES IN AUTOMATIC PROCESSING OF SUBSTANCE CUES

The impetus for this work has come from theoretical models that emphasise the importance of automatic cognitive processes as an essential mediating step that can account for the effects of substance-related cues on craving and substance-related behavior. It has been argued that substance users find it difficult to ignore substance-related cues in their environment, which exacerbates subjective craving in response to those cues, and this in turn leads to substance-seeking behavior. Additionally, attentional bias for substance cues may occur as a consequence of a dysfunctional, dopamine mediated incentive learning process, which causes substance-related cues to grab the attention and elicit automatic approach responses, ultimately resulting in increased craving and substance-seeking behavior.

Attentional bias refers to the observation that substance-related cues can "grab" or "hold" the selective attention of substance users. Studies using the addiction Stroop task have demonstrated that users of alcohol, tobacco, cannabis, cocaine, and heroin are slow to name the color in which substance-related words are presented and this impairment increases across levels of substance usage. This suggests that substance abusers find it difficult to disregard task-irrelevant alcohol-related cues, which leads to an impairment in the primary task (color naming). Other studies have used the visual dot probe task, often combined with eye movement monitoring, which provides more direct measures of visuo-spatial attention. Results obtained from this task demonstrate that substance abusers (including heavy drinkers, heroin users, and tobacco smokers) are faster to respond to probes that appear in the location of substance-related pictures compared to probes that appear in the location of control pictures. These populations also maintain their gaze for longer on substance-related pictures, which suggests that substance abusers bias their spatial attention toward the location of the substance cues. However, some nicotine research has found attentional bias, as measured by the visual dot probe task, to be greater in lighter, compared with heavier, smokers. It is possible that this reflects a shift from goal-directed to more habit-like behavior, indicating that attentional bias plays a less important role in heavier substance use.

More sophisticated measures of biases in selective attention have emerged in recent years, such as the P300 component of event-related brain potentials, and these have also been applied to demonstrate attentional biases in addiction. Although the P300 is often found to be blunted in response to general stimuli in substance-dependent populations, P300 amplitude has found to be greater in reaction to substance-specific cues. In addition, P300 amplitude has been found to positively correlate with self-reported substance craving and, therefore, may demonstrate the motivational significance of substance cues.

Another cognitive process which may be important in the development of substance use disorders is "appraisal," that is, how a person appraises the importance of interoceptive or exteroceptive cues in relation to current behavior and goals. Some cognitive models suggest that as addiction progresses, appraisal biases develop which result in information being processed in accord with an individual's goals (e.g. substance taking). Appraisal is likely to be affected by attentional biases and "interpretative biases," whereby ambiguous situations or states (e.g. physiological arousal) will be interpreted within a substance-relevant context. Research shows that ambiguous stimuli are more likely to be interpreted as alcohol-related in those who drink more and score higher on measures of alcohol use disorders. Some cognitive models presume that it is the attention to, and the appraisal and interpretation of the cue that is crucial in cue reactivity, rather than the cue itself. This proposal may explain why some research fails to

find congruent relationships between traditional measures of cue reactivity and craving. For example, if cognitive processes are focused on the exteroceptive cue but not on the internal autonomic state, then physiological measures of reactivity will be minimal and not relate to self-reported craving. Individual differences may exist concerning what type of cue effect (e.g. behavior, affective, physiological) is focused on and so, if this is accepted, cue reactivity may actually be diverse across individuals and levels of substance use disorder. More research is needed to investigate these possibilities, however, it does highlight a theoretical shift from physiological and subjective responses to substance-related cues (as reviewed in the previous sections), to subtle biases in the cognitive processing of those cues.

At present, there is no evidence to suggest that attentional biases operate below the threshold of conscious awareness, although results obtained using the alcohol Stroop task suggest that distraction that is created by alcohol-related cues occurs automatically, and is difficult to control or impede. However, it is interesting to note that in visual dot probe task studies, heavy drinkers tend to show attentional biases for alcohol stimuli only when they are presented for relatively long exposure durations (500 ms or more) but not when they are presented briefly (e.g. 200 ms). As such, attentional bias among heavy drinkers who are not seeking treatment may primarily reflect a bias in the maintenance of attention, or delayed disengagement of attention from alcohol-related cues. By contrast, in alcohol-dependent inpatients (compared to nonalcoholic controls), attentional biases are seen for briefly-presented (50 ms) alcohol-related stimuli, and the magnitude of this effect is related to the severity of alcohol dependence. Also, among alcohol-dependent inpatients, if the stimuli are presented for long periods of time, attentional avoidance of those stimuli is seen. The approach-avoidance pattern of attentional bias that is observed among treatment-seeking alcoholics may reflect motivational conflict or ambivalence: the initial orienting may reflect sensitization of the incentive value of alcohol, whereas the subsequent avoidance may occur because alcohol-related cues are aversive when presented in a treatment context. Overall, the relationship between craving and attentional bias tends to be quite weak, however, research suggests that the association between subjective craving and the latter component of attentional bias (delayed disengagement of attention) is stronger than the relationship with the earlier component (rapid initial orienting). Given this, the approach-avoidance pattern of attentional bias among patients in treatment may actually reflect the aversive properties of alcohol cues in the treatment context (leading to rapid initial orienting toward alcohol cues) coupled with diminished subjective craving (leading to a diminution or even reversal of the bias to maintain attention on alcohol-related cues).

Substance-related cues also elicit automatic approach tendencies in substance abusers. For example, in the stimulus–response compatibility task, participants are required to rapidly categorise substance-related and neutral pictures by moving a manikin either toward or away from the pictures as quickly as possible. In one block of the task, participants are required to make the manikin approach substance-related pictures and avoid neutral pictures; in a different block of the task, these instructions are reversed. Substance abusers (but not control participants) are faster when required to approach, rather than avoid substance-related pictures, and this has been demonstrated in heavy drinkers, tobacco smokers, and cannabis users. The mechanisms that underlie these effects are presently a matter of debate: substance-related cues may trigger an automatic approach response as would be predicted by the cognitive urge and automaticity model discussed earlier, which results in participants being faster when required to respond to those cues by performing a (motivationally compatible) approach response rather than an (motivationally incompatible) avoidance response. An alternative explanation is that these results reflect strengthened underlying associations between the concepts of "substance" and "approach" in substance abusers, therefore one might not expect substance-related cues to elicit an automatic approach tendency unless the substance-relatedness of those cues was explicitly encoded. Nonetheless, experimental manipulation of automatic response tendencies elicited by alcohol-related cues can have a causal influence on drinking behavior in the laboratory, which suggests that these automatic approach responses may have a causal effect on substance-seeking behavior.

The research on cognitive biases in addiction has helped demonstrate the complexities of cue reactivity and highlight how biases may change across contexts (e.g. in treatment and nontreatment environments) and, therefore, how cue reactivity may also differ. Factors which impact on how cue reactivity is translated into behavior need to be identified by future experimental research and applied to clinical populations and treatment interventions.

NEUROBIOLOGY OF CUE REACTIVITY

Recently, functional neuroimaging studies have revealed the neural substrates of cue reactivity and associated cognitive biases. These studies have identified a number of brain areas which are involved in cue reactivity, including the mesolimbic dopaminergic pathways (reward, motivation) as well as the amygdala (memory),

anterior cingulate cortex (attention), dorsolateral prefrontal cortex (inhibitory control), orbitofrontal cortex (reward, motivation, decision making), hippocampus (memory), and thalamus (memory and attention).

Meta-analysis shows that there is considerable overlap in the neural substrates for cue reactivity to the most studied substances (i.e. nicotine-, alcohol- and cocaine-related cue reactivity). Cue reactivity activates the ventral striatum, anterior cingulate cortex, and the left pallidum. Importantly, these are also the areas that seem to be involved in self-reported craving. Although the search for the neural substrates of the cognitive bias associated with cue reactivity – using Stroop-related tasks – has just begun, it seems that the dorsal anterior cingulate cortex, the superior parietal gyrus, and the superior temporal gyrus play a role. These are brain regions associated with top–down processing, suggesting that substance-dependent individuals have to employ more attentional resources to focus on alternative tasks when substance cues are present.

Functional magnetic resonance imaging research shows that non-dependent student populations, with some evidence of substance problems, respond to substance cues differently to neutral cues. The similarities between these findings with those from dependent populations suggest that neural-based cue reactivity may develop very quickly as hazardous substance use progresses, or that such reactivity may reflect a pre-existing risk for substance problems. In addition, some research has found associations between cue-induced brain activation and craving only in moderate, not severely, dependent populations. This may indicate a shift from goal-directed substance behavior to habit-like behavior, the latter perhaps being more likely to occur with or without cues.

Although the majority of research suggests a complex system of neural activity in response to cues, which integrates areas involved in incentive salience, attention, memory, information processing, decision-making, and goal-directed behavior, there remain discrepancies within the literature. One issue is that the perceived availability of a substance differs across cue reactivity studies, and this may be one determinant of inconsistent results. For instance, if the participant expects to smoke after completing a nicotine cue reactivity task, this will likely affect processes of expectancy and preparedness which will be apparent in neural activity. Treatment-seeking dependent populations are trying to resist substance use, and are using their cognitive resources to resist relapse, so results must be interpreted in light of the current goals of the participant (e.g. treatment-seeking, not treatment-seeking). Another issue is the specific substance under study. Although there are many common neural substrates involved in cue reactivity across substances, there are also some substance-specific differences. All these factors will impact on the nature of the cue reactivity observed and needs to be taken into account when identifying the clinical application of such research.

CLINICAL APPLICATIONS

Perhaps the most researched cognitive and behavioral treatment to have emerged from the cue reactivity literature is cue exposure during extinction; by removing the reward associated with a conditioned stimulus or an operant response, an organism eventually stops expecting, or responding for, the reward. The effect of extinction has been found to reduce reactivity to a wide range of substance cues, however, rather than the original association being unlearned, a number of observed phenomena suggest that the original learning persists but that new learning also takes place. Firstly, the conditioned response can be "reinstated" if the unconditioned stimulus is encountered following extinction. "Renewal" of the conditioned response can occur when the context differs from that of the extinction environment. For example, if acquisition (e.g. context A) and extinction (e.g. context B) of an association occur in different environments, then renewal may be observed if the organism re-enters context A (ABA renewal effect). In fact, renewal of the conditioned response may be seen in any environment that differs from the extinction context (ABC renewal effect). The conditioned response can also show "spontaneous recovery" when the organism comes into contact with a cue after a prolonged period of time. Finally, organisms will learn extinguished associations more quickly than new associations, indicating that the original association is still intact (i.e. "rapid reacquisition"). In its current form, cue exposure treatment does not seem to have a significantly beneficial effect on relapse rates. Taken together, these findings suggest that to increase the efficacy of cue exposure treatment, exposure sessions would need to be conducted across a wide range of different environments and for a prolonged period of time, to help ensure the new learning can be generalised and strong enough to compete with the original cue–substance associations.

Currently, "talking therapies" are the standard psychosocial treatment for addiction, and these include cognitive behavioral therapy, motivational interviewing, and group therapy. When comparing standard treatment with standard plus cue exposure treatment, research tends to find that the inclusion of cue exposure increases the patient's self-efficacy regarding their ability to resist substances in high-risk contexts. However, even when a decrease is observed in cue

reactivity following exposure treatment, there is not always a decrease in craving or relapse and some research finds higher relapse rates following cue exposure treatment. In addition, a reduction in cue reactivity has been observed irrespective of treatment type, therefore, reduced cue reactivity may be a general effect of treatment procedures.

Given the behavioral methods needed to increase the generalisation and strength of new associations developed during treatment, the cost-effectiveness of successful cue exposure treatment is questionable. This has resulted in research aimed at identifying pharmacological agents which may enhance cue exposure treatment. D-cycloserine (d-4-amino-3-isoxazolidinone) is a glycine partial agonist which activates the N-methyl D-aspartate (NMDA) receptor complex, a type of ionotropic glutamate receptor. The NMDA-glutamate synapses are crucial in learning and memory, and NMDA antagonists can block learning, therefore, exciting these connections during learning may strengthen new associations. The effects of D-cycloserine on learning diminish as the time between extinction procedures and D-cycloserine administration increases, suggesting that D-cycloserine is facilitating acquisition and/or consolidation of the new association. In addition, animal models have shown some generalisability of effects. If animals learn two conditioned associations and complete extinction trials to just one of these, administration of D-cycloserine can also reduce responding to the nonextinguished conditioned stimulus. This finding indicates that extinction-based addiction treatments used in conjunction with D-cycloserine may result in more general extinction effects which may, in turn, reduce the types of environment in which cues can precipitate substance use. However, other work has found the effects of D-cycloserine to be context-specific and more recent cue reactivity studies in humans have failed to find a beneficial effect of D-cycloserine. Although it is unclear whether treatments which focus on the learned associations in addiction are likely to result in significant decreases in intake levels and relapse, it is apparent that the key to making such treatments efficacious is to make their effects generalise to multiple contexts.

Other pharmacological agents which may offer therapeutic benefit via effects on cue reactivity processes include naltrexone, a μ-opioid receptor antagonist. Treatment with naltrexone, compared with placebo, can reduce self-reported urge following substance cue exposure. Other main pharmacotherapies for alcoholism include acamprosate (which has some of its effects on the gamma-aminobutyric acid system) and disulfiram (inhibits alcohol metabolism), however, findings are equivocal in terms of the impact these drugs have on cue reactivity. Recently, the efficacy of drug treatments has been linked, at least partly, to the genetic profile of the individual. This research is in its infancy and it remains to be determined which functional genetic variants are associated with different pathways into substance problems (e.g. via cue reactivity, impaired inhibitory control). However, some research has indicated that an amino acid substitution (Asn40 to Asp40), caused by the rs1799971 polymorphism on exon 1 of the OPRM1 gene, is associated with cue reactivity. The Asp40 variant leads to a μ-receptor subtype with up to three times greater binding of β-endorphin than the more common Asn40 variant, and craving in response to alcohol cue exposure is more pronounced in carriers of the Asp40 variant. In addition, evidence suggests that naltrexone may have more efficacious general treatment effects in carriers of the Asp40 variant. Although an association between the Asp40 variant and cue reactivity has yet to be fully established, the research illustrates, in principle, that genotype might modulate cue reactivity and thereby treatment response.

Models which see cue reactivity as a more unique response pattern, difference across individuals, suggest that the cognitive mechanisms (e.g. appraisal, beliefs, expectancies, biases) held by the individual need to be mapped so that interventions can be more targeted. Although in general, reducing attentional and approach biases to substance cues may provide some therapeutic benefit, understanding the individual differences which affect the development and nature of such biases should improve the efficacy of any bias-based treatments. Some studies suggest that attentional bias has a causal role in alcohol drinking, and some positive treatment results have been found, suggesting attentional bias retraining may be a suitable target for treatment development. However, other work has failed to replicate early successes and has called into question the causal role of biases in substance use.

SUMMARY

Cue reactivity covers a wide range of subjective, physiological, and neurochemical responses, and can be observed in social, hazardous, and dependent substance populations. The state changes that occur during cue exposure are likely to contribute to initiation and maintenance of substance use in a significant number of people, however, cue reactivity is neither necessary nor sufficient to account for the chronic relapsing nature of addiction. Perhaps one important finding is that cues appear to have an effect on behavior which is relatively autonomous from the actual incentive value of the substance. Such findings should inform treatments aimed at educating people about the dangers and negative consequences of substance use.

Although the early conditioning and learning research within this area has been invaluable, it is clear that cue reactivity is complex and highly individual, involving both substance-like and substance-opposite effects. More recent research highlights the potentially important role that cognitive biases may play in cue reactivity, and identifying how and why these development offers a promising area for future research which may help better predict cue reactivity across different populations and environments.

Given the potential importance of cue reactivity, treatments based on cue exposure have delivered disappointing results. Future work may benefit from developing exposure treatments which also include cognitive bias retraining, ensuring that the various processes involved in reactivity are recognised. The possibility that pharmacological and cue exposure treatments may work synergistically is important. However, the growing awareness of how individual factors (e.g. genetics) affect treatment response indicates that some kind of screening process is required. Screening treatment-seeking individuals should help target treatments and increase their cost-effectiveness, whether these are treatments based upon the evidence from cue reactivity research or not.

SEE ALSO

Alcohol Use Disorders, Heroin Addiction, Cocaine Addiction, Marijuana Use and Abuse, Methamphetamine Addiction, Tobacco, Binge Drinking, Cognitive Factors in Addictive Processes, Contextual Factors in Addiction, Relapse and Lapse, Tolerance and Withdrawal, Implicit and Associative Processes in Addiction, Sensory Imagery in Craving, Deprivation, Craving, and Affect: Intersecting Constructs in Addiction, Craving and Expectancies, Relation of Craving and Appetitive Behavior, Defining and Assessing Drug Craving, Attentional Biases in Craving

Glossary

Approach bias an automatic tendency to approach substance-related stimuli.

Attentional bias an automatic tendency to direct attention toward substance-related stimuli.

Automaticity the process by which behavior is elicited in the absence of intention to perform the behavior. Automatized behavior tends to be stimulus bound and difficult to inhibit.

Classical/pavlovian conditioning a learning process in which a previously neutral stimulus (conditioned stimulus) becomes associated with a biologically significant stimulus (unconditioned stimulus) which naturally elicits a response (unconditioned response). As a result, the conditioned stimulus comes to produce a response (conditioned response) by itself.

Cue exposure treatment a treatment approach in which patients are exposed to cues associated with their addiction (e.g. a glass of wine, a wrap of cocaine) in the absence of substance administration. The aim of cue exposure treatment is to extinguish associations between the cues and substance effects.

Cue reactivity a phenomenon in which exposure to substance cues produces a range of physiological (e.g. alterations in heart rate, respiration, and temperature) and psychological (e.g. substance-related expectations and substance-relevant cognitive biases) responses, which motivates the individual to seek out and administer substances. Cue reactivity is believed to play a role in initiating and maintaining substance use and may be an antecedent for relapse.

Negative reinforcement the process in which a behavior is more likely to be repeated if that behavior leads to reduction or removal of a negative affective state.

Operant/Instrumental conditioning a learning process in which an association is learned between a response and an outcome. Cues can act as a discriminative stimulus and trigger the response, either through direct stimulus–response or stimulus–outcome(expectancy)–response associations.

Pavlovian-to-instrumental transfer (PIT) a learning process involving two phases: a classical/pavlovian conditioning stage, in which a conditioned stimulus (e.g. a cue) is associated with an unconditioned stimulus, and an instrumental conditioning stage where an association is made between a response and an outcome which matches the unconditioned stimulus. Following these two stages, the conditioned stimulus (cue) can elicit the response for the outcome (unconditioned stimulus).

Positive reinforcement the process in which a behavior is more likely to be repeated if that behavior leads to an increase in positive effect.

Further Reading

Carter, B.L., Tiffany, S.T., 1999. Meta-analysis of cue-reactivity in addiction research. Addiction 94 (3), 327–340.

Conklin, C.A., Tiffany, S.T., 2002. Applying extinction research and theory to cue-exposure addiction treatments. Addiction 97 (2), 155–167.

Drummond, D.C., 2000. What does cue-reactivity have to offer clinical research? Addiction 95 (Suppl. 2), S129–S144.

Drummond, D.C., Tiffany, S.T., Glautier, S., Remington, B., 1995. Addictive Behaviours: Cue Exposure, Theory and Practice. Wiley & Sons Ltd, UK.

Field, M., Cox, W.M., 2008. Attentional bias in addictive behaviors: a review of its development, causes, and consequences. Drug & Alcohol Dependence 97 (1–2), 1–20.

Field, M., Munafo, M.R., Franken, I.H., 2009. A meta-analytic investigation of the relationship between attentional bias and subjective craving in substance abuse. Psychol Bull 135 (4), 589–607.

Franken, I.H.A., 2003. Drug craving and addiction: integrating psychological and neuropsychopharmacological approaches. Progress in Neuro-Psychopharmacology and Biological Psychiatry 27 (4), 563–579.

Hogarth, L., Duka, T., 2006. Human nicotine conditioning requires explicit contingency knowledge: is addictive behaviour cognitively mediated? Psychopharmacology (Berl) 184 (3–4), 553–566.

Kühn, S., Gallinat, J., 2011. Common biology of craving across legal and illegal drugs – a quantitative meta-analysis of cue-reactivity brain response. European Journal of Neuroscience 33 (7), 1318–1326.

Perkins, K.A., 2009. Does smoking cue-induced craving tell us anything important about nicotine dependence? Addiction 104 (10), 1610–1616.

Ryan, F., 2002. Detected, selected, and sometimes neglected: cognitive processing of cues in addiction. Experimental and Clinical Psychopharmacology 10 (2), 67–76.

See, R.E., 2002. Neural substrates of conditioned-cued relapse to drug-seeking behavior. Pharmacology Biochemistry & Behavior 71 (3), 517–529.

Stewart, J., de Wit, H., Eikelboom, R., 1984. Role of unconditioned and conditioned drug effects in the self-administration of opiates and stimulants. Psychological Review 91, 251–268.

Tiffany, S.T., 1990. A cognitive model of drug urges and drug-use behavior: role of automatic and nonautomatic processes. Psychological Review 97 (2), 147–168.

Relevant Websites

www.niaaa.nih.gov – National Institute on Alcohol Abuse and Alcoholism.

www.nida.nih.gov – National Institute on Drug Abuse.

www.ias.org.uk – The Institute of Alcohol Studies.

CHAPTER

44

Craving and Expectancies

Ross McD. Young

Institute of Health and Biomedical Innovation, Queensland University of Technology, QLD, Australia

Many will try a range of substances in adolescence or early adulthood, but only a small proportion will go on to develop problematic use. As such, the experience of a substance misuse problem is not something that most individuals can identify with personally. However, the experience of craving an object or activity and the importance of making predictions about the expected consequences of our future behavior are concepts that have universal significance. While we accept the use of the term craving in everyday conversation, there has been a considerable scientific debate about the semantics of craving. Regarding addictive behavior, there is a general agreement that craving reflects a desire or urge to consume a substance – to either experience pleasure or provide relief from discomfort. This aligns with clinical experience, where those with drug dependence commonly report a sensation they call craving that parallels this definition. On this basis alone, craving is worthy of consideration in models of addiction.

The wish or urge to consume a substance that we refer to as craving is a broad cognitive-emotional construct. Craving is acknowledged in many models of alcohol and other drug use as contributing to the development and maintenance of drug-related problems. As such it features in some definitions of dependence, including that of the International Classification of Diseases, ICD-10. However, craving is not experienced by all drug users, and even those with severe dependence can report an absence of craving. Disentangling individual differences that may contribute to craving is therefore important. Craving is shaped via learning processes, but there are also important neurobiological factors that have a primary influence on craving state.

Craving is acquired via direct experience with a drug as a consequence of prior consumption. The stronger intensity of craving experienced for some drugs may also be a feature of their use and be instrumental in maintaining addiction to the drug, for example in the case of heroin or nicotine dependence, where craving and associated physical symptoms may be marked. In contrast, physical symptoms are not as prominent in stimulant withdrawal. It can be difficult to differentiate intense craving from the emergence of the withdrawal state, and while both may be related, the withdrawal state typically has a predictable time course and emergence of a complex array of symptoms. Craving may be one of these. The withdrawal syndrome, or elements of withdrawal, is not simply a consequence of declining drug concentrations, excessive CNS receptor stimulation, and tolerance, but is also related to conditioning mechanisms that pair environmental cues with drug administration. Thus, craving can be viewed as a phenomenon that while independent of withdrawal, is a component of the subjective experience of withdrawal related to drug deprivation for many drug users.

The desire for a drug does not exist in the absence of consideration of the perceived consequences of drug ingestion. These drug outcomes are more complex than simply the relief of withdrawal or a general sense of drug reward. Indeed, both desire to use and anticipation of reinforcing outcomes related to use could be present in individuals where no withdrawal symptoms are experienced. Consideration of these subjective drug outcomes has included exploration of motives, attitudes, intentions, and expectancies. Outcome expectancies (or expectations) reported by drug users are acquired as a result of both indirect and direct experiences of drug use. Unlike craving, drug expectancies can be present before the first use of a substance. These expectancies represent a more finely grained cognitive construct than the cognitive–affective construct of craving, as they relate to an array of specific outcomes and consequences that may occur once a drug has been taken. Drug expectancies are typically described as substance-specific, for example as alcohol

http://dx.doi.org/10.1016/B978-0-12-398336-7.00044-9

expectancies or cannabis expectancies, as various drugs have different outcomes due to their pharmacological properties, typical modes of use and environments of administration. Given that these expectancies link anticipated ingestion of a drug with a subsequent outcome, they are viewed as learned *if–then* contingencies. For example, *if* I were to drink now, *then* I would feel sadness (a negative expectancy); or *if* I were to smoke a cigarette now, *then* I would feel more relaxed (a positive expectancy). Expectancies have relevance in the spectrum of use from prior to first consumption to dependence. Prior to first use expectancies are acquired via observation, modeling, and vicarious reinforcement. In dependent use, they are potentially based on decades of direct experience with a substance. For example, expectancies of alcohol as a social lubricant or reliever of stress have been identified in children as young as 5 years of age, and coalesce with actual drinking experience in adolescence. The nature of the drug expectancies an individual holds is predictive of subsequent drinking behavior patterns. Tension reduction expectancies are markers of risky consumption of several drugs including alcohol, cannabis, and heroin. Once acquired via direct experience with the drug, this memory network of positive expectancies can be primed by internal or external drug-associated cues. Primed expectancies are thought to guide subsequent drug use.

Drug expectancies are not the only cognitions that are likely to shape use, or be relevant in a consideration of craving. Efficacy expectations that relate to the expected confidence with which a substance might be refused in some future circumstance are also important. In a similar way to drug outcome expectancies, there are several domains of drug refusal self-efficacy, and they have also been identified in adolescents with minimal experience with alcohol. Drug refusal self-efficacy expectancies can impact consumption decisions, future use, and the nature of problems associated with consumption. There is a good evidence that drug expectancies and refusal self-efficacy operate in concert, and that both an individual's drug expectancies and confidence in refusal are probable casual factors in drug consumption. Most of the data regarding drug expectancies are from studies of alcohol expectancies: there is a stronger body of work across several drug classes supporting the importance of self-efficacy beliefs.

Triggers of craving include expectancy, with some models viewing this memory of the rewarding effects of use combining with an underlying negative affective state to generate drug urge. Urges are likely to ulitize attentional capacity, making expectancies and other associated cognitive outputs such as images more difficult to shift from working memory. These cues can be internally generated via expectancy once established and can create a powerful motivational state to use.

Craving, once considered an epiphenomenon, is now along with drug expectancy seen as an important aspect of drug motivation. Despite craving being commonly reported by those with drug dependence as a core feature of their experience, the emphasis on craving in models of addiction has been in and out of fashion over the last 60 years. A key debate has been around whether craving represents a meaningful causal factor in the development of addiction. The challenges of measurement have also been a key impediment to progress. The disease model of the 1950s and 1960s emphasized craving as a physical manifestation of the disease of addiction. Behavioral models, which emerged in the 1970s, saw a reconceptualization of craving, but there was still considerable disagreement regarding issues of definition and the threshold when desire for a substance was strong enough to be considered a craving state. There was little consensus as to whether or not the concept was of value. However, a commonly held definition of craving relating to the "subjective desire for the effects of a drug" views it as a potential summary index of potential reinforcement.

The renaissance of craving as a construct has spawned several theories since the 1980s. Some have posited craving as a consequence of unconscious cognitive processes, which only come into awareness when the automatic process of drug intake in those with established dependence is disrupted. This emphasizes cognitive processes where introspection is limited, or not possible at all. Others, based on classical conditioning processes, view craving as a consequence of seeking to alleviate drug withdrawal and clearly consider that self-reports of craving are reliable given that the drug is negatively reinforcing. In many ways these are a psychological extension of the earlier medical model, where aversive consequences were emphasized, but in addition they examine the contribution of learning to withdrawal-related craving. These models, unlike those of earlier medical models, do not view craving as "physiologically inevitable." Environmental cues associated with drug use can, through conditioned association, elicit a significant withdrawal response. For example, a heroin user can, by being simply present in an environment such as a "shooting gallery" where they typically inject, experience situational withdrawal discomfort including physical symptoms and a longing for the drug.

Subsets of conditioning-based models of craving describe homeostatic processes that exert their effect over time as tolerance to the appetitive effects of the drug. These effects occur along with the emergence of compensatory responses. The compensatory responses are opposite to the effects of the drug (e.g. if the drug is anxiolytic, or anxiety relieving, then the compensatory response would be anxiogenic, or anxiety inducing).

These compensatory responses are hypothesized to be the basis of withdrawal and contribute to craving for withdrawal relief. The relative balance of positive and negative expectancy would be a likely reflection of these dynamic changes over time, particularly for such striking subjective effects as a shift from anxiolytic to anxiogenic outcomes. If models emphasizing these opponent processes were ubiquitous, it would be unlikely that positive expectancies would remain in those with significant dependence, as the compensatory responses would have predominated over the appetitive effects and the experience of use would become aversive. Severely dependent users typically still report positive expectations of use, although this is in the face of increasingly strong negative expectancies.

A contrary view to models based on withdrawal relief is that environmental cues are associated with drug administration as signals of reward. This has led to two sets of models regarding craving, which emphasize conditioned cues as either signals of anticipated positive reinforcement (also called craving type-1) or signals of withdrawal (also called craving type-2). There is evidence to support both types of model: both types of craving coexist and may vary between drugs and level of dependence.

The theories that underpin craving clearly influence, and are influenced by, the measurement of the phenomenon of craving itself. Given the approaches previously mentioned, which are predicated on withdrawal relief or reward, there are measures that conform to each of these two types of craving, but most measure both. Simple single item measures and multifactorial scales exist, and the more sophisticated measures based on withdrawal and reward models do not necessarily have better utility in predicting dependence or relapse. Most of the available drug-specific measures are of alcohol craving. Scale content may include obsessive and intrusive thoughts about alcohol, the intensity of the compulsive urge to drink and sensory aspects of craving. Typically these measures are summarized by two to four factors, but replication of underlying factor structures has been a challenge, limiting the selection of a preferred definition or measure. Despite these measurement challenges, across the body of measures in use, craving is consistently associated with dependence severity and poorer prognosis. Craving exists in nondependent drinkers, and while multifactorial scale structures have been replicated in those with less developed or absent problems, the predictive utility of craving measures under these circumstances is restricted. Whether this is due to scales being less sensitive to craving of lower intensity or whether there are qualitative differences in craving in those who are not drug dependent compared to those with dependence, is unclear. Additional challenges are evident in tapping the more subtle manifestations of emergent craving, as measured by self-report questionnaires in those without established dependence. These measures have stronger validity in those with drug dependence. In dependent users, imagery may be a particularly important component of craving, with stronger craving being associated with more intense sensory imagery, for example of the sight, taste, or smell of alcohol.

Given the role of positive and negative expectancies as reflections of the motivation to use drugs, it is likely that both reward and relief craving are present at different stages of the evolution of dependence to a varying degree, with relief predominating in cases of more severe dependence. Cognitive models of craving have emphasized higher order learning that is both potentially positively and negatively reinforcing and are thus of direct relevance to a consideration of drug expectancies. Indirect evidence exists to support the role of changes in cognitive processing in a craving state which may prime expectancies. For example, smokers in withdrawal focus more on smoking-related stimuli and on the positive consequences of use in their nicotine-deprived state, indicating that both craving and expectancy are linked.

The potential to integrate our understanding of craving and expectancy has been advanced by better knowledge of the brain processes that underlie addiction. Psychobiological addiction research has expanded our notions of learning related to drug use by explicating the underlying physiological mechanisms of reward. Contemporary advances in molecular biology and neuroimaging have driven this. This emergent understanding forms a sound basis to examine potential links between craving and expectancy learning.

There is strong evidence that craving is influenced by several neurotransmitters including dopamine, gamma-aminobutyric acid (GABA, the primary inhibitory neurotransmitter), endogenous opioids, glutamate, and serotonin. All drugs of abuse, except the benzodiazepines, impair dopamine reuptake, enhance dopamine release, or reduce dopamine breakdown. For example, cocaine blocks the active reuptake of dopamine via the dopamine transporter in the shell of the nucleus accumbens (NAcc), a structure in the ventral striatum implicated in reward prediction and adaptive behavior expression. This impaired reuptake increases dopamine availability in the synaptic cleft resulting in stronger signals of reward. Brainstem neurons rich in dopamine D2 receptors (DRD2), particularly in the ventral tegmental area (VTA), have a strong role in approach and avoidance learning and the mediation of reward. These receptors are stimulated by drug ingestion via enhanced dopamine release. This dopamine system has been extensively investigated in the development

of addictive behavior. Projections extend from the VTA to the NAcc and orbitofrontal cortex (OFC), and collectively this is often referred to as the dopamine "reward pathway." GABAergic interneurons (which are stimulated by benzodiazepines and thus indirectly influence dopamine release in the VTA) are embedded within midbrain dopamine neurons and this region also contains localized glutamatergic neurocircuitry. Collectively this reward pathway integrates inputs from several key structures involved in sensory processing, memory, motivation, and planning (e.g. from the amygdala, which responds to the intensity of cues related to reward or punishment) and it projects to higher order cognitive centers (e.g. OFC) to promote associative learning. This reward pathway and the neurotransmitters that influence it provide a plausible link between the affective component of craving and the conscious cognitions of expectancy, given the signals of incentive salience generated.

Animal studies have confirmed the importance of the neurotrophin, the brain-derived neurotrophic factor, and the secreted protein glial cell-line derived neurotrophic factor, in the synaptic and structural plasticity of mesolimbic dopamine neurons. This operates via the production of drug-related experience-dependent dendritic spine proliferation. The richness of these connections is in essence a marker of previous drug-related learning built via operant conditioning and represents learned associations. The enriched connective network of the mesocorticolimbic dopamine system and its effect on the reward pathway enhances the "natural" function of this system of performance monitoring based on external feedback. Importantly, dopamine is released to the drug-associated cue rather than the reward itself.

Conditioned place preference studies provide a useful experimental paradigm to examine operant learning. They measure the relative probability that animals that have been administered drugs via operant conditioning paradigms return to environments where the drug was administered, as opposed to an alternative environment where drugs have not been administered, when they are given a choice. This place preference has been examined using dopamine agonist drugs administered during operant trials (where learning is enhanced) and dopamine antagonist drugs (where learning is impaired). Other studies have used rodents that have been specifically genetically bred with an absence of DRD2 (D2 knockouts) and this drug experience dependent learning is completely absent. These studies support the mediation of crucial learning processes related to possible if–then contingencies and desire via dopamine release. The potential for drugs of abuse to stimulate brain pathways that have evolved to facilitate approach and avoidance reward learning to enhance survival, as well as to create dynamic changes in synaptic connectivity based on drug use, is a core component in the development of addiction. For example, stronger amphetamine-induced dopamine release in midbrain structures in humans has been examined via neuroimaging and is associated with increased drug-related "wanting" and novelty seeking. Importantly, this increased dopamine release is correlated with more positive subjective ratings of the effects of amphetamine. Neuroendocrine human and animal studies indicate the complexity of these processes extend beyond the neurotransmitters noted and involve endocrine effects where leptin, ghrelin, and thyroid stimulating hormone (peptides and hormones involved in food seeking) are also associated with drug seeking or acute craving. This observation occurs particularly in human studies examining those with an early onset of dependence. Collectively, these results further support the notion that processes that subsume learning about natural reinforcers related to improved survival via effective reward learning are the underlying key elements of drug reinforcement.

Examining these processes in more detail has commonly involved the manipulation of brain reward systems pharmacologically, and examining individual differences in response. Bromocriptine, a specific DRD2 agonist used in the treatment of Parkinson's Disease, a disorder on diminished dopamine function, boosts reward learning and related responses in the striatum of those with the A1+ allelic version (Taq 1A) of the DRD2 gene (also called the Ankyrin repeat and kinase containing 1 or ANKK1 gene). No similar enhancement occurs in A1− individuals who do not possess this allele. The DRD2 agonist drug cabergoline has produced similar striatal responses. A1+ carriers also have a significantly lower concentration of DRD2 in several midbrain structures including the NAcc and striatum. They also have fewer postsynaptic D2 autoreceptors in the striatum. These autoreceptors act to dampen dopamine release, so when they have diminished function related to this genetic marker, the likely result is a greater tonic release of dopamine from dopamine agonist drugs of abuse. A1+ allelic status is probably a broadly based risk factor, rather than specific for any drug or class of drugs, as individuals with this allele are more impulsive and show more drug-related physical and psychological comorbidity due to their impulsivity. For example, A1+ individuals who abuse heroin and are intravenous drug users are more likely to be infected with Hepatitis C, presumably due to a lessened delay of gratification despite sterile injecting equipment being unavailable.

Psychobiological models also illuminate the affective element of craving. Disruptions in midbrain neurotransmitter systems related to chronic drug use, particularly

alterations in NAcc dopamine and serotonin activity, and increased corticotropin releasing factor in the central nucleus of the amygdala, contribute to a negative affective response. This is a key component of craving and it may also influence expectancy development. There is an increased firing of subgroups of dopamine neurons in the VTA in response to stress, which may constitute a stimulus for further drug seeking. Indeed, the importance of dopamine in craving and reducing negative effect has been confirmed via placebo-controlled study using bromocriptine. Craving for alcohol was reduced in those receiving the active drug, but only if they were A1+ in terms of DRD2 gene status. A similar effect was evident for anxiety reduction, but over a different time course, indicating that craving (as measured by a single item craving scale) and anxiety could be differentiated. Clearly both were influenced by dopamine. No effect was evident for those with A1− status or those administered placebo.

Both translational and reverse translational research have extended our knowledge of drug reinforcement and craving. There is a strong emphasis on translational research in the delivery of treatment for physical and psychological disorders to ensure that the benefits of theoretical findings are incorporated into the delivery of more effective and efficient care. The psychobiological research already discussed is beginning to produce better outcomes, particularly through the development of medications that alter the systems that give rise to drug reward and craving. Other evidence to support underlying neurobiological processes that contribute to addiction comes from the trials of medications used to treat addiction themselves.

The role of GABA has been confirmed in several studies of baclofen, a $GABA_B$ receptor agonist, which results in decreases in craving, anxiety, and alcohol consumption. Baclofen also reduces dopamine release in the dorsal and ventral striatum in animal experiments.

Mu-opioid receptors are associated with reward and are present in the VTA. These VTA receptors when activated by endogenous opiates or opiate drugs of abuse inhibit GABA, thus reducing their inhibitory effect on dopamine. Increases in mu-opioid receptors over time are associated with alcohol and cocaine craving. Naltrexone, a mu-opioid receptor antagonist, is an effective treatment for alcohol dependence when combined with psychosocial treatment and appears to operate via reductions in craving when exposed to alcohol-related cues. This may then make those seeking treatment more receptive to skills acquisition including cognitive skills aimed at reducing positive expectancies. Those receiving naltrexone also reported less interest in drinking and a reduction in alcohol-related "highs" if they lapsed during the course of treatment, than did those on placebo. This prevents the escalation of a lapse to a relapse.

Glutamate, the most common excitatory neurotransmitter in the brain, stimulates the glutamatergic N-methyl-D-aspartate (NMDA) receptor and projections from the prefrontal cortex, amygdala, and hippocampus to the NAcc release dopamine, including indirectly via the VTA. In a similar manner to the increased dendritic spines on dopamine VTA neurons that are responsible for conditioned association learning during cued drug administration, the richness of these glutamatergic connections with the NAcc with dopaminergic afferents from the VTA increase as a product of previous drug exposure. This is a thought to contribute to drug-seeking behavior. Acamprosate acts at the NMDA receptor and enhances GABAergic inhibition. There is some evidence that acamprosate reduces craving for alcohol but this has not been consistently replicated. It may act via reducing excitatory responses to alcohol-related cues.

The outcomes of psychological intervention trials have also been illuminating in terms of refining our understanding of the role of desire and subjective outcomes. Psychological treatments based on exposure theory provide repeated unreinforced exposure to drug-related cues predicated as the primary means of reducing craving. There is evidence that exposure trials of at least 15 min are necessary for benefit. However, differential effects between cue exposure and other effective treatments such as cognitive behavior therapy (CBT) have not been evident. Typically habituation within exposure sessions and across trials has been demonstrated, but this has also been evident from CBT, suggesting that the reduction of craving due to environmental exposure may be common to multiple psychological therapies. Alternative treatments based on models contrary to the principles of cue exposure can also reduce craving. Mindfulness-based relapse prevention involves the teaching of skills to observe emotional and cognitive states such as craving through relaxation and other exercises without reacting in the habitual manner. The goal is to be more aware of responses rather than to extinguish them. Mindfulness appears to break the association between negative effect and craving. Following usual care incorporating 12-step and psychoeducational approaches, craving and negative affective states are still related after completion of therapy. In mindfulness-based therapy, repeated unreinforced exposure to the internal negative affective state, rather than to an external cue, may be an important advance in more effectively creating new learned associations.

Drug expectancies reflect conditioned associations and drug reward mechanisms. Expectancy studies have mapped the potential domains of reinforcement and have examined underlying processes associated with dependence such as salience, tolerance, and craving. As previously noted, positive expectancies are

consistently related, in their number and strength, to reports of drug consumption. The brain processes underlying learning and the acquisition of addiction have direct relevance to the development of drug expectancies. For example, $GABA_A$ receptor beta-3 subunit allele status is associated with alcohol expectancy related to negative effect, linking GABA-related genetic influences with self-reported alcohol outcomes. This genetic influence extends beyond the central nervous system to the effectiveness of alcohol metabolism and outcome expectancies. ALDH2*2 status confers protection against alcohol dependence through an aversive response to alcohol via the accumulation of acetaldehyde. Acetaldehyde is a toxic metabolic product of alcohol that causes increased heart rate, flushing, and nausea and in most individuals is rapidly metabolized. Presence of the ALDH2*2 allele is associated with lessened positive expectancies about alcohol, again indicating a relationship between individual differences in physiological mechanisms related to drug response and beliefs about drinking outcomes.

It is also assumed that drug expectancies are conditioned to cues because of their temporal link with drug consumption and are subsequently reinforced by drug ingestion itself, and this temporal effect is likely to be more distal from the act of drug use than craving itself. Primed expectancies prior to drug consumption are also likely to elicit or strengthen the acute experience of craving. In one study, cross-sectional data indicated that positive opiate expectancies are predictive of reward craving in opiate-dependent individuals, but a similar effect for alcohol expectancy was not evident for alcohol dependence. By contrast, a lack of positive coping strategies was predictive of high relief craving in opiate dependence, and high reward craving in alcohol dependence.

The in vivo study of drug effects offers some key benefits related to experimental control. Analog studies involving the administration of alcohol to those without alcohol dependence typically use convenience samples of undergraduate students. They do not support a strong role for craving but typically support the influence of alcohol expectancies on in vivo drinking behavior.

Manipulation of positive expectancies to decrease expectancy intensity through laboratory or therapeutic procedures shows a casual relationship with lowered consumption. For ethical and practical reasons these studies have typically focused on alcohol. Procedures designed to decrease expectancies are called expectancy challenges (EC). EC manipulations of positive expectancies have usually taken place in simulated bars, with college or undergraduate students. Most often, participants have been young adults with a short drinking history of moderate to heavy alcohol consumption rather than dependence.

In an EC procedure, a socially pleasant atmosphere is generated in the laboratory bar, and drinks are made available to participants. In the tradition of other alcohol research using placebo-controlled designs, half of the participants are given an alcoholic beverage, and the other half are given a placebo, which tastes and smells of alcohol but has no alcoholic content. In some studies all beverages have been placebo. The socially pleasant atmosphere is designed on the basis of these environmental cues to elicit both positive alcohol expectancies and craving.

In an early EC two sessions were separated by 5 days, where male participants were administered either alcohol or a placebo. In the first session, word games were played and turns were taken through the group with the session designed to facilitate social interaction. The second session had a sociosexual theme. Photographs of female models from magazine advertisements were displayed one by one, and discussions of their relative attractiveness took place to arrive at a consensus judgment. Toward the end of these sessions, participants were asked to guess who had consumed drinks containing alcohol and who had consumed a placebo beverage. Mistaken judgments were used to demonstrate that expecting that alcohol had been consumed could cause an individual to behave as though alcohol had actually been consumed. The importance of alcohol expectancies in generating what was otherwise considered to be an alcohol-related pharmacological response alone was the theme of the discussion. The goal was to disrupt the association between drinking and expectancy. In the final session, these initial observations were summarized, and the role of alcohol expectancies in driving future consumption was emphasized. As a consequence of the EC, heavy drinkers reported weaker endorsement of sexual and social alcohol expectancies and reduced their alcohol consumption. No measures of craving were taken, but based on the nature of the population, craving was likely to have been weak. Changes in positive expectancies and consumption were not evident in a control group exposed to a standard college prevention program instead of the three EC sessions, or in a control group that completed only before and after measures. Moderate drinkers exposed to EC reported no change in consumption. Subsequent studies have replicated the effect for social and sexual expectancies, but changes in expectancies related to affective and cognitive arousal did not occur. Equally, there has been replication for males but not for females, unless they are heavy drinkers. Using more traditional methods of CBT can also diminish the strength of the positive expectancies and increase the strength of negative expectancies, particularly those related to social enhancement.

The motivational systems that shape drug use include cue-related information, expectancies, and signals of

drug availability. In addition to altering expectancies to change this motivational system, it is viable to target urges in those with the most risky expectancies. Naltrexone, the mu-opioid antagonist used as relapse prevention agent, has been studied in this regard. Naltrexone can diminish the urge to drink in hazardous drinkers who do not have manifest dependence, but only in those with high positive expectancies. In other words the maximum benefit of naltrexone appears to be when positive expectancies are primed. The treatment implications of this finding in those with established dependence have yet to be explored.

Taken together, these findings show that positive alcohol expectancies can be reduced through EC in those with alcohol abuse. However, even in the absence of expectancy change, given the motivational network that incorporates craving and expectancy, reducing the urge to drink in those with strongly positive expectancies via effective pharmacotherapy shows promise.

Drug refusal self-efficacy, drug expectancies, and craving collectively shape drug use motivation. A lapse is likely to reflect multiple influences ranging from craving to impulsive decision making. The notion that the final decision was caused by a lack of belief in personal agency is common. For example, "I didn't have the confidence to refuse an offer of cannabis and just gave in to temptation." Social learning theory describes the acquisition and maintenance of all human behavior, including alcohol use, as derived from two related but independently operating expectations: outcome expectancies, such as alcohol expectancies which have already been described and self-efficacy expectancies. Self-efficacy expectancies reflect an individual's belief in their capacity to regulate a specific behavior, and these expectancies reflect the person's situational confidence. They may influence effort in the face of a challenge directly or interact with the person's mood state: they may, for instance, both induce and be undermined by depression.

There are self-efficacy measures for several drugs of misuse, with some important common elements across different classes. The broad nature of self-efficacy domains identified via factor analysis is almost identical across different drugs of abuse, despite their different pharmacological actions. A common factor in alcohol, cannabis, heroin, and nicotine self-efficacy involves difficulties in resisting substances in situations where a person needs emotional relief. Challenges to self-efficacy in social pressure situations are also evident with alcohol, nicotine, and cannabis users. Drinking refusal self-efficacy has been associated with adolescent and university student drinking patterns in cross-sectional studies. Cross-sectional data from adults seeking help have confirmed that self-efficacy is more strongly associated with the number of occasions of alcohol consumption than with the quantity consumed on a specific occasion, indicating that these beliefs may be more important in the initiation of substance use on any given occasion relating to alcohol-related cues.

There is good evidence for the influence on the development of self-efficacy beliefs, from learning related to the reward pathway. For example, those who have DRD2 gene A1+ status show diminished drinking refusal self-efficacy in situations of social pressure, compared with their A1− counterparts. Importantly, those with A1+ allelic status report an earlier age of onset of alcohol-related problems, which is in turn associated with lower drinking refusal self-efficacy and subsequent level of alcohol consumption and dependence. The association of A1+ status with lower drinking refusal self-efficacy, but not with greater alcohol expectancy, underscores the importance of these reward systems in shaping "wanting" to drink as opposed to "liking." "Wanting" is likely to driven by the cue-conditioned pulsatile dopamine release in the VTA. Furthermore, using the "reward value" of alcohol (how rewarding do you expect it to be to have alcohol presented to you today) as a summary index of "wanting" explains additional statistical variance in craving over and above prior drinking behavior and personality.

Despite the more limited predictive utility of craving in those without misuse, research that examines the association of craving and expectancy and involves the presentation of alcohol tends to be conducted in convenience samples without significant drug problems. As previously discussed, this is for ethical and practical reasons and usually involves college or undergraduate students. Some of this work has confirmed an association between the ratings of valence, dominance, and effect associated with alcohol-related cues and of craving with alcohol expectancy strength. Other studies have failed to find a relationship between expectancy strength and reactivity to alcohol-related cues.

Substance abuse treatment research has historically focussed more on "what treatments work" rather than "why they work." There have been repeated claims since the 1980s that effective treatment increases patients' self-efficacy. Self-efficacy expectancies are likely to develop through past experiences; enactive mastery; vicarious learning from peers, parents, social reference groups, and the media; and emotional experiences. The least effective influence on self-efficacy is verbal persuasion. Confirming the mechanisms responsible for clinical improvement has proved elusive, and mechanisms to date have typically been inferred rather than definitively identified. There is a large body of work indicating that CBT, based on social learning principles, is effective. CBT specifically attempts to bolster self-efficacy through teaching effective problem solving

skills, rehearsal, and skills to challenge dysfunctional expectancies. CBT fares no better than other bona fide treatments when direct comparisons are made. Thus, techniques designed to enhance social skills, challenge maladaptive cognitions, and teach drinking refusal skills do not show comparative benefit over techniques that do not purport to operate via cognitive-behavioral mechanisms. CBT studies have not identified reliable behavioral or cognitive mediators of treatment outcome, including self-efficacy.

When alcohol expectancy and self-efficacy are examined together, self-efficacy beliefs appear to be stronger predictors than expectancy of alcohol consumption at posttreatment. However, some studies have failed to find evidence for a predictive role of self-efficacy. Self-efficacy predicts relapse to heavy drinking over the first few months post abstinence, and is typically a more effective predictor of posttreatment relapse than expectancies over periods ranging from 6 months to 2 years. Reductions in positive alcohol expectancy and increases in self-efficacy have been demonstrated in a handful of studies, but they have not concurrently examined craving. Self-efficacy has predictive power for heavy drinking over 12 months in those who are taking the anti-craving drug *naltrexone* over and above effects related to craving reduction.

Craving itself can be elicited by alcohol expectancies, and imagery scripts based on positive alcohol expectancies have been used as an experimental paradigm to investigate craving, where urge and strength of personally held alcohol expectancies were related. The strength of craving in alcohol-dependent individuals from the experimental use of imaginal cues, based on autobiographical memories of craving, is as strong as those from actual exposure. Craving and expectancy may work in concert to enable behavior change under these circumstances. For example, effective CBT has been associated with expectancy change, particularly related to social and relationship relevant domains. Abstinence across 12 weeks of CBT is associated with reductions in assertion and sexual enhancement expectancy and increases in social pressure self-efficacy. Extinction of craving may be enhanced via exposure to social situations where drinking was previously undertaken as a consequence of increased confidence to resist alcohol in these environments.

These processes are not fixed: the priming of expectancies and craving requires information of drug availability. For example, the craving experience in drug-dependent individuals within inpatient settings is significantly less intense than in those who are continuing to use drugs in their usual settings. While there are multiple confounds acting on this relationship, there is a good evidence from nicotine dependence studies that craving for nicotine varies as a function of being told that there is a potential opportunity to smoke a cigarette versus being told that smoking is not possible. The impact of exposure to smoking-related cues is attenuated when smokers believe that there is little opportunity to smoke in the near future as compared with being informed that smoking would be possible soon. Cues may elicit craving via expectancies that include the broad expectancy of the possibility to access substances. These processes are not independent of the reward processes already described. There is brain imaging evidence that greater OFC activation in substance users who considered that the opportunity to use was imminent, compared with those for whom use was not possible and were trying to remain abstinent. In smokers who were not attempting to cease smoking, a subsequent opportunity to smoke during cue exposure resulted in greater brain activation, including in several subregions of OFC, than the activation in those who were told that cigarettes were not available. However, there was no association between this instructional expectancy set and craving. In a similar manner, these instructional expectancies regarding availability partially mediate attentional bias to smoking-related cues. The generation of drug expectancies may also involve particular subregions of frontal cortex. Again on the basis of neuroimaging, individual decision making under risk is negatively associated with dorsomedial prefrontal cortex (DMPC) activation and positively associated with ventromedial prefrontal cortex (VMPC). Those with strong activation of VMPC are more sensitive to reward on that basis display a higher probability of risky decision making.

SUMMARY

The motivation to use drugs is influenced both by the beliefs we hold about the consequences of drug ingestion and by the urge to use. Drug expectancies in conjunction with an array of other internal and external cues can elicit craving. Craving is a cognitive-emotional state of desire to use drugs that many dependent drug users experience, and it can relate to both the avoidance of an aversive state and seeking reward. Both types of craving coexist and are potentially related to drug expectancies although the experience of craving is more closely temporally linked to drug use itself. Individual differences in the neurobiology of reward related to genetics and learning history influence both craving and expectancy development. Drug expectancy, refusal self-efficacy, and craving can all be stimulated by external drug-related cues in the presence of signals of drug availability.

Craving has fallen in and out of fashion, but has been given more credence over recent years as an important

component in models of drug use, given its neurobiological validation. Although not directly compared in research studies, drug expectancies appear to develop first via observational learning and then both craving and expectancy develop as a consequence of use itself. Both craving and expectancy are shaped via brain processes that signal the presence of reward and facilitate associative learning. While these processes are complex and involve multiple neurotransmitters, dopamine has a central role. Drug expectancies and drinking refusal self-efficacy can work together as part of the same memory network to influence the acute initiation of drug use. There is a growing body of work confirming viable ways of altering expectancy and craving. In the absence of expectancy challenge, diminishing motivation via craving reduction in the context of strong expectancies may be a future useful treatment. The treatment implications of integrating expectancy and craving are yet to be fully realized.

SEE ALSO

The Biopsychosocial Model of Addiction, Cognitive Factors in Addictive Processes, Relapse and Lapse, Cue Reactivity, Implicit and Associative Processes in Addiction, Sensory Imagery in Craving, Interference with Concurrent Tasks, Deprivation, Craving, and Affect: Intersecting Constructs in Addiction, Neural Correlates of Craving for Psychoactive Drugs

List of Abbreviations

CBT cognitive behavior therapy
DRD2 dopamine D2 receptors
EC expectancy challenges
GABA gamma-aminobutyric acid
NAcc nucleus accumbens
NMDA N-methyl-D-aspartate
OFC orbitofrontal cortex
VMPC ventromedial prefrontal cortex
VTA ventral tegmental area

Glossary

Allele a variant, or form, of a gene.
Autoreceptor a presynaptic receptor on a neuron that is sensitive to neurotransmitters produced by that neuron, providing a feedback loop that controls neurotransmitter release.
Drug expectancy a belief about the likely behavioral, cognitive or emotional effect of a drug on the person.
Drug refusal self-efficacy confidence in being able to refuse drugs that are offered in specific circumstances (e.g. under social pressure).
Glial cells cells that surround and protect neurons, providing oxygen and nutrients. Their role in neurotransmission is still being explicated.
Mesocorticolimbic pathway a dopamine pathway from the midbrain to the frontal cortex, that has been implicated in the salience of incentives or rewards.
Neurotrophins growth factor proteins that modulate the functioning and life cycle of neurons.

Further Reading

Drobes, D.J., Carter, A.C., Goldman, M.S., 2009. Alcohol expectancies and reactivity to alcohol-related and affective cues. Experimental and Clinical Psychopharmacology 17, 1–9.

Heilig, M., Thorsell, A., Sommer, W.H., et al., 2010. Translating the neuroscience of alcoholism into clinical treatments: from blocking the buzz to curing the blues. Neuroscience and Biobehavioral Reviews 35, 334–344.

Goldman, M.S., 2002. Expectancy and risk for alcoholism: the unfortunate exploitation of a fundamental characteristic in neurobehavioral adaptation. Alcoholism: Clinical and Experimental Research 26, 737–746.

Kavanagh, D.J., Andrade, J., May, J., 2005. Imaginary relish and exquisite torture; the elaborated intrusion theory of desire. Psychological Review 112, 446–467.

Leggio, L., 2009. Understanding and treating alcohol craving and dependence: Recent pharmacological and neuroendocrinological findings. Alcohol and Alcoholism 44, 341–352.

MacKillop, J., Monti, P., 2007. Advances in the scientific study of craving for alcohol and tobacco. In: Miller, P., Kavanagh, D. (Eds.), Translation of Addictions Science into Practice. Elsevier, Oxford, pp. 189–209.

Marlatt, G.A., 1985. Cognitive factors in the relapse process. In: Marlatt, G.A., Gordon, J.R. (Eds.), Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors. Guilford Press, New York, pp. 128–200.

Palfai, T., Davidson, D., Swift, R., 1999. Influence of naltrexone on cue-elicited craving amongst hazardous drinkers: the moderational role of positive alcohol expectancies. Experimental and Clinical Psychopharmacology 7, 266–273.

Robinson, T.E., Berridge, K.C., 1993. The neural basis of drug craving: an incentive-sensitization theory of addiction. Brain Research Reviews 18, 247–291.

Skinner, M.D., Aubin, H.-J., 2010. Craving's place in addiction theory: contributions of the major models. Neuroscience and Biobehavioral Reviews 34, 606–623.

Volkow, N.D., Li, T.K., 2004. Drug addiction: the neurobiology of behaviour gone awry. Nature Reviews Neuroscience 5, 963–970.

Young, R. Mc.D, Lawford, B.R., Nutting, A., Noble, E.P., 2004. Advances in molecular genetics and the prevention of substance misuse: Implications of association studies of the A1 allele of the D2 dopamine receptor gene. Addictive Behaviors 29, 1275–1294.

Young, R. Mc.D, Oei, T.P.S., 1993. Grape expectations: the role of alcohol expectancies in the understanding and treatment of problem drinking. International Journal of Psychology 28, 337–364.

CHAPTER

45

Attentional Biases in Craving

Jon May

School of Psychology, University of Plymouth, Drake Circus, Plymouth, UK

OUTLINE

ATTENTIONAL BIASES IN CRAVING

In everyday terms, an attentional bias might be interpreted simply as a tendency to attend more toward some thing than others, and it would thus seem straightforward to conclude that people who are dependent upon a substance will pay more attention toward it, and notice it more in their environment, particularly if they are currently craving it. Within cognitive psychology, however, facilitated recognition or responding to a cue is described as priming, and can arise for many different reasons, including prior exposure, motivation, overlearning, or general activation of mental representations (*see* Cognitive Factors in Addictive Processes by Neighbors). In this context, attentional bias has a more precise meaning, since it reflects a bias within the attentional processes themselves, and it has been argued that this may represent a change in mental functioning that causes or maintains substance dependence. If this is the case, then therapeutic approaches that modify such biases may help alleviate craving, and hence assist abstinence attempts. Theories that see addictive behavior as an automatic response (such as Tiffany's view that drug-seeking behavior is driven by the activation of overlearnt action schemas, with craving occurring when these schemas are blocked or inhibited) would predict a strong relationship between attentional bias and craving, and also between attentional bias and drug use.

An alternative possibility is that attentional biases do not play a causal role in craving, either being a consequence of craving, or being caused by the same processes that lead to craving, and so while they may attenuate as craving is reduced, attempts to modify them directly will not have any direct impact upon craving, substance use, or addiction. For example, Robinson & Berridge's Incentive Sensitization hypothesis suggests that use of addictive substances leads to neurological changes in the brain making users hypersensitive to drug-related cues (*see* Incentive Salience and the Transition to Addiction by Robinson, Berridge, & Robinson). While these changes might lead to both an increased desire for the drug, and an increased tendency to attend to drug cues, modifying the attentional bias would not be expected to reverse the drug-related neurological changes, even if it had behavioral effects in the form of reduced addictive behavior or reports of subjective craving.

A more cognitive interpretation is that an attentional bias to substance-related cues might lead to increased conscious awareness of the substance, and hence a subjective state of craving that distracts the individual from other tasks, motivating drug-seeking behavior. This link between automatic and controlled cognition is a key component of, for example, the Elaborated Intrusion Theory of desire proposed by Kavanagh, Andrade, and May (EI Theory). Where other theories of addiction emphasize the automatic or bottom-up control of

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00045-0

drug-seeking behavior, EI Theory gives conscious strategic behavior a mediating role. Automatic processes, including attentional biases, lead to intrusive thoughts about substance use, but craving only occurs if the individual cognitively elaborates the initial thought, especially through substance-related sensory imagery (*see* Sensory Imagery in Craving by Andrade). Conscious thoughts of substance use may in turn moderate attentional allocation, confounding attempts to measure purely automatic attentional biases in deprived substance users or those manipulated into states of high craving. While EI Theory predicts a relationship between attentional bias and craving, the link is one of many, and is not direct or one-way.

A better understanding of the relationship between addiction, craving, and attentional biases would be very useful in identifying the relative contributions that explicit, conscious processes (such as craving) and implicit, automatic processes (such as attentional biases) play in addiction (for a detailed account of implicit processing, *see* Implicit and Associative Processes in Addiction by Wiers and Stacy). One difficulty in untangling the causal relationships is that many attentional bias studies using substance-related materials do not measure craving, or do not measure it at the same time as the attentional bias measure. Often they simply show differences between dependent and non-dependent groups. Other studies may compare substance-dependent participants who have or have not been deprived, or who have been exposed to cues relevant to a substance. Increasingly, the role of craving is being recognized, and short scales that simply ask participants to rate how strongly they are experiencing an urge to use or consume a substance are administered several times during studies, allowing baseline craving and changes during the study to be assessed. In a recent review of such studies of visual attentional bias, Field, Munafò, and Franken concluded that while a correlation between visual attentional biases and craving did exist, it was weaker than might be expected from theories that give attention a direct role in driving drug-seeking behavior.

Historical Aspects

Attentional bias was initially described and developed theoretically within the context of affective disorders, and in particular anxiety, by Andrew Mathews, Colin MacLeod, and colleagues at St George's Hospital in London. Their work had focused upon the role that negative schemas of thought might play in anxiety. It had long been recognized that depressed people tended to interpret the world around them in a negative manner, giving negative interpretations to everyday events that other people might see as neutral or even positive. When asked to recall events, they would include more negative material or events than non-depressed people, and would tend to dwell on or ruminate on these thoughts. Early work on attentional biases evolved as an attempt to explain differences in the patterns of findings between depressed and anxious people.

In the 1960s, Aaron Beck had proposed a cognitive model of affective disorders that focused upon the negative content of thoughts, in contrast to the then-dominant behaviorist model that saw emotional problems as a set of learned responses to stressful or threatening situations. The model was criticized for lacking an explanation of how these negative thoughts came about or were processed: without such an explanation, the model was difficult to test empirically. Using the then-dominant cognitive concepts, Gordon Bower proposed an associative network-based model of mood and memory, in which emotional states were represented as nodes within a general semantic memory network, alongside other nodes representing events and memories. When one memory or node was activated, it would in turn activate any nodes to which it was connected or associated. In this model, affective disorders resulted from a richer linking of negative emotional nodes to other memories, such that almost every thought or stimulus would spread activation to negative emotions.

This model allowed empirical, hypothesis driven research to take place which confirmed the predictions of memory biases in depression, especially explicit biases, where people were asked to search memory for answers. For anxiety, however, the results were different. Instead of explicit memory biases, it seemed that implicit memory biases were typically being found.

Taken together with anecdotal evidence from clinicians about anxiety patients being hypervigilant for threat cues in their environment, it seemed that instead of a memory bias, anxiety might be linked to an attentional bias. Mark Williams recounted the cases of a bird phobic who was unable to walk down her high street in case the butcher's shop there had a display of poultry in the window, and of a welder who obsessively scanned through newspapers looking for reports of disasters. For these people, their elevated levels of stress and anxiety could be seen as a natural and correct response to the highly threatening world that they found themselves in, for the more they searched the world for instances of the things they feared, the more they found them.

Stroop Task

Initial studies on attentional bias used two methodologies taken from experimental psychology, one based on Stroop's color-name interference effect, and the other on

dividing auditory attention in dichotic listening tasks. Stroop had asked people to name the color of the ink that had been used to print various color names. When the color name and the ink matched, people were fast to do this task; but when the color name was different to the ink, reading speed slowed and errors increased, due to the tendency people had to read out the color name instead of naming the ink.

The Emotional Stroop was a variant upon this in that emotional words were printed in different colors. Gordon Bower reported an experiment that he had carried out with Gerald Clore in which participants who had been induced into an emotional state were slower to name the ink color of emotional phrases than neutral phrases, whereas participants in a neutral mood named the colors of all phrases equally and rapidly. Fraser Watts and his colleagues adapted the test to show that spider phobics were slower to color-name words associated with spiders, and C. Ray found that students were slower to color-name words related to examination anxiety than neutral words, when tested just before their exams.

To see if this effect was specific to anxiety or if it was involved in other motivational states, other researchers looked at the effect of food deprivation on color-naming food-related words, and found similar results. By using substance-related words instead of emotional words, a task known as the Addiction Stroop was developed. In one of the first reports of an Addiction Stroop, Gross and colleagues showed that smokers who had abstained overnight were slower to color-name smoking-related words than neutral words, while non-abstinent smokers showed the opposite pattern. This finding is important because it does not just distinguish between dependent and non-dependent participants, but shows that the level of deprivation is associated with the interference effect.

The Addiction Stroop is attractive because it is easy to use. In its simplest form it can be prepared as a printed sheet of around 30 words, and the time taken to color name the entire sheet recorded, and divided by the number of words used (see Fig. 45.1). The effect is then the difference in time per word obtained for a sheet of substance-related words and a set of neutral words – although to control for category effects the neutral words should ideally all be drawn from a coherent set, such as furniture or transport words. More complex versions of the task can be conducted using computer-based presentation, in which words can be presented individually for precise durations, allowing masking to be used (neutral stimuli presented before or after the cue). Stimuli can be presented in blocks or mixed with neutral stimuli.

However, the task is controversial and difficult to interpret. Slower naming speeds on any Stroop test could be due to the difficulty of inhibiting processing

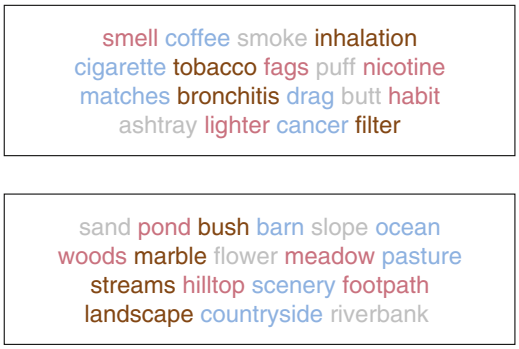


FIGURE 45.1 Addiction Stroop task – an example of an Addiction Stroop task using sets of 18 smoking-related and environmental words. Participants name the ink color of each word (pink, gray, blue, or brown), and the effect is measured by the difference between the mean naming time for the two sets. Smokers craving cigarettes would take longer to color name the smoking words than the environmental words. *Word sets used by Waters, A.J. and Feyerabend, C., 2000. Determinants and effects of attentional bias in smokers.* Psychology of Addictive Behaviors, 14(2), *111–120.*

of the word content, as originally suggested, or due to a gradual build up of associated thoughts due to the use of repeated salient words from a single category. Even if the effect is due to individual words, one or two verbal stumbles in a set of 30 words are sufficient to create an overall difference in color-naming time. It is not clear that the test is measuring anything about attention at all, and even if it is, it is not a pure test.

In a comprehensive meta-analytic review of the use of Addiction Stroop tests, Cox, Fadardi, and Pothos found 18 studies using alcohol cues, 11 smoking cues, 3 drug abuse cues, and 1 gambling cues. While they found that variations in the method of presentation affected the sensitivity of the task to addiction-related attentional bias, the task did reliably show such an effect, and the effect was increased by manipulations that induced or elicited addiction-related thoughts. Increases in the amount of use, or the desire to use, substances also increased the effect. However, they also noted that no study had explicitly measured craving at the same time as administering the Stroop task, although one study reported correlations between heroin craving in the prior week and interference on a heroin-Stroop task. It seems that the task is sensitive to the increased activation of substance-related concepts, and while it seems highly plausible that this is also related to craving, the link is not yet proven.

Cox, Fadardi, and Pothos make a number of recommendations for studies using Addiction Stroop tasks. Four colors should be used, and in computer-based tasks, requiring participants to say the color name out aloud in addition to pressing response keys enhances the effect. Stimulus words need to be chosen carefully so that there are no associations with color, and the neutral words should form categories that really are

neutral, with no emotionally charged content. The addiction and neutral categories should contain an equal number of words, and should be balanced for frequency and length. Blocked presentation is also preferred over intermixed presentations in computer-based testing, but if intermixed presentation is to be used both the addiction and neutral words need to be preceded by additional neutral filler words to reduce the chance of effects from one trial persisting into the next trial (which is presumably a factor in the increased sensitivity of blocked presentation). They also recommend that very short and very long tasks should be avoided, and although they make no recommendation about the appropriate number of trials per category, the median number of words per category in the reviewed studies was 12, with each word being used four times, for a total of 48 trials per category.

Dual Task Procedures

Another, more obviously attentional, early approach to measuring attentional biases used a dual task procedure known as dichotic listening, in which different audio streams are played to each ear over headphones. The basic design requires people to perform some primary task upon the stream played in one ear, the 'attended' channel, and then tries to assess how much they are distracted by events played in the other 'unattended' ear. This can be measured by a decrease in performance upon the primary task, or by an increased ability to detect the 'unattended' events.

Dual task paradigms are widely used in experimental psychology to study the degree to which different mental faculties are independent of one another (if the two tasks do not interfere), or load upon shared resources (if they do interfere). In the context of attention research, the assumption that attention cannot be divided but must switch between tasks, sources of information, or spatial locations means that an attentional bias can be detected if performance on the primary attended task is reduced more by addiction-related content in the unattended channel than by neutral content.

Mathews and MacLeod adapted the dichotic listening task and asked anxious and non-anxious control participants to shadow (i.e. to repeat while listening) eight stories played to one ear, while ignoring word lists read in the other ear. Some of these lists contained threatening words and others were entirely neutral. They hypothesized that the anxious participants would have their attention drawn toward the unattended threat lists, but not to the neutral lists, and to measure this participants were also asked to complete a simple reaction time task by pressing a response button whenever a cue appeared on a computer screen. As predicted, the response times for the non-anxious participants were unaffected by the word list content, at around 500 ms, whereas the anxious participants took 600 ms to respond during the neutral word lists and 680 ms during the threatening word lists. Following the task, participants all reported being unaware of the unattended word lists, and memory tests for the words used in the unattended lists indicated no recall. This indicates that the attentional bias toward the threatening words was preconscious and not a result of increasing awareness during the task.

In an addiction context, Sayette and Hufford found that manual responses to tones were slower when deprived and non-deprived smokers were holding and looking at a lit cigarette in their other hand (at around 350 ms) than if they were holding and looking at a similarly sized spool of electrical tape (300 ms). They also measured self-reported craving, and found that in the group of smokers who had been deprived of cigarettes and then exposed to cigarette cues, there was a significant relationship between craving and the amount of slowing in response times; non-abstinent smokers and abstinent smokers who had not been exposed to cigarette cues did not show this association. This suggests that the attentional bias was due to the cigarette exposure, and not to deprivation, but that when smokers were also deprived, the strength of their craving was associated with an increased attentional bias. The craving seems to enhance an attentional bias brought about by awareness of cigarette cues, rather than the cues inducing an attentional bias that leads to craving.

As with the Stroop effect, it is not clear that dual tasks are a purely attentional effect, because the repeated presentation of negative or threatening material might have a cumulative effect upon mood, which might in turn affect responses or performance on a primary task. Similarly, presenting cues related to substance use might trigger associated thoughts about that substance that could also distract a dependent participant and affect their performance, without the cues themselves being selectively attended to.

Visual Probe Tasks

Andrew Mathews and Colin MacLeod turned to the attentional literature to search for a purer measure of attention, and adapted a visual cueing technique developed by Michael Posner. In what has come to be called the dot probe or visual probe task, two cues are presented simultaneously in different parts of a computer screen (see Fig. 45.2). When the cues disappear, one of them may be replaced by a small probe or target, often a dot, and the participant has to respond to this probe. There are many ways on which the task can vary, as described

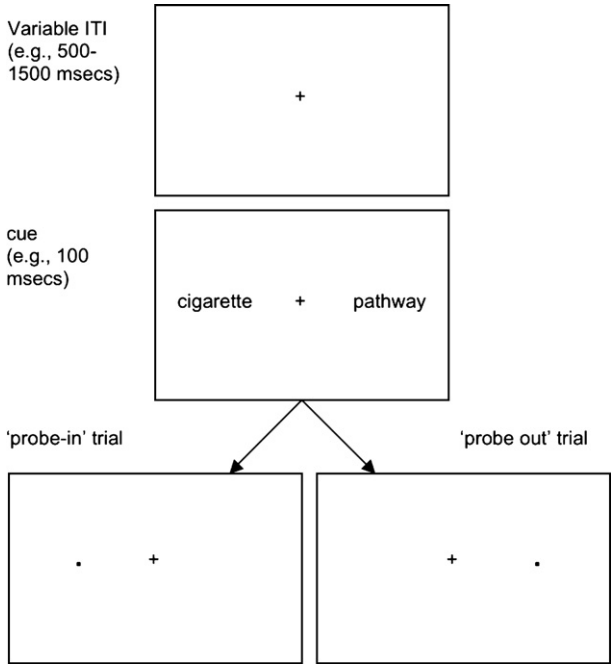


FIGURE 45.2 Visual probe task – a trial from a typical visual probe task, using smoking-related and neutral stimuli. The participant is asked to look at the fixation cross, and to press one response button when they see a dot appear on the left of the screen, and a different button if it appears on the right. Deprived and craving smokers' responses are faster on 'probe in' trials when the probe letter is displayed in the location of a prior cue that is smoking related than on 'probe out' trials, when it follows the neutral cue, indicating an attentional bias toward the smoking cues. The opposite pattern (avoidance of smoking cues) may be found for non-smokers.

below, but a key feature is the simultaneous onset of two stimuli which thus compete for attentional resources. If one of them is especially salient in some way (e.g. a threat word for anxious people, or a substance-related word for users of that substance) then participants' attention may be drawn toward the location of that cue. Response latencies to probes that appear in this location will therefore tend to be faster than those presented in the location of the other cue. The difference in latencies between the two locations gives an index of attentional bias.

The basic task is quite flexible in that it can use verbal or pictorial cues, which can be separated horizontally or vertically by differing amounts, and presented either subliminally (i.e. for less than 50 ms), very briefly but supraliminally (between 50 and 500 ms), or for much longer periods which allow the participant to attend to, look at, and then look away from the cue. Participants can be asked to fixate on a central location before the cues appear, and not to look at them, or be allowed to move their gaze. The probe (and critical cues) can appear on all trials, or on a subset; and the probe can require a simple detection response or (more usually) a binary decision, contingent upon probe location or probe content. These variations are all interesting from a theoretical analysis of the functional location of the effect within the cognitive system: but the presence of the effect is robust – if at times small, being measured in milliseconds.

Mathews and Macleod's first study reporting the use of this task found that anxious participants responded 45 ms faster to threat location probes than non-threat location probes, whereas the non-anxious participants were 24 ms slower. In fact, the tendency of the supposedly neutral control group to show an avoidant, rather than a neutral, pattern of responding, is a frequent finding in anxiety studies, which suggests that normal behavior involves avoidance of threat. Over the last 25 years, variants of the visual probe task have been widely used in studies of affective disorders, especially anxiety, but also in other motivational contexts, including addiction.

The use of two cues in the visual probe task is based upon the idea that attention can only be focused upon one spatial location (or stream of information) at a time, and so using two simultaneous cues forces the two cues to compete for attention. Sensitivity to any differences in attentional bias toward the experimental cue is increased, whereas the onset of a single cue always attracts attention. A related procedure in which only one cue is presented at a time has also been used, and is called the attentional cueing task (see Fig. 45.3). Because attention is inevitably allocated toward the

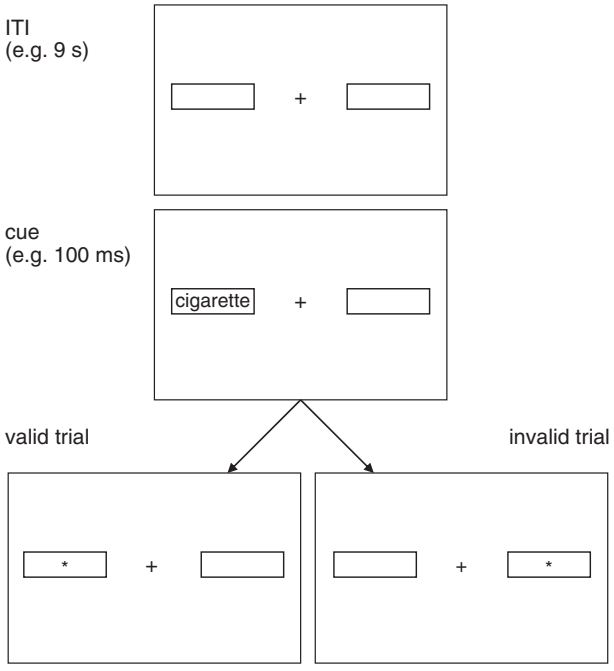


FIGURE 45.3 Attentional cueing task – a trial from a typical attentional cueing task. Unlike the visual probe task, only one cue is presented at a time. Participants are asked to respond as quickly as possible when they detect a probe, and substance users show slower response on invalid than on valid trials. Note the longer inter-trial interval (ITI) required compared to the visual probe task.

single cue, it is more sensitive to differences in disengagement of attention from the cue, or post-attentional avoidance.

Field and Cox recently reviewed the findings from attentional cueing and visual probe studies on users of alcohol, cannabis, cocaine, heroin, and tobacco. While they found that in general, substance users tended to show an attentional bias toward their substance, two patterns stood out. First, subliminal cues, which were not reportable by participants and so not available for conscious processing, did not lead to attentional bias. Secondly, alcoholics who were undergoing treatment showed a pattern of avoidance, with slower detection of probes in the location of alcohol cues. This would seem to indicate that, as with the Stroop task, attentional bias effects in addiction may be due to explicit, conscious processing rather than very early, automatic processing: this is the opposite to the findings from the anxiety literature.

New Methodological Developments

As the attentional literature evolves, tasks developed there can be adapted to measure attentional biases in more sophisticated ways. For example, some researchers are now using the attentional blink phenomenon in which the ability to detect a briefly presented target within a sequence of distractors is impaired if it closely follows a distractor that is salient in some way (typically, being visually similar to the target). While this phenomenon has been extensively studied using varying visual stimuli to explore visual attention, Barnard and colleagues used words as stimuli to show that increasing the semantic relatedness between distractor words and the target category could also lead to attentional blink effects, showing that the effect is not purely caused by visual effects but that meaning is also important. Other researchers have also shown that emotionally valenced targets are less prone to the effect, which suggests that they are more able to regain attention following a distraction event. Addiction researchers have now used the task to show that smokers show smaller attentional blinks for smoking-related targets than for non-smoking targets. Although these studies have not yet assessed different levels of craving, they suggest that this task might be another useful way of measuring the degree and timeframe of attentional biases.

Other researchers are making increased use of eye-tracking tasks, in which shifts in gaze direction are measured. While gaze direction follows the allocation of attention, and so can be used as an indirect measure, it is not identical to attention. Studies using eye tracking, however, have shown that substance users gaze for longer at substance-related cues, as if they find it hard to move their attention away from them. These techniques can usefully be combined with visual probe tasks, to separate trials involving explicit changes in gaze direction from those that do not: if attentional biases can still be found in the absence of changes in gaze direction, then relatively early components of attention would be contributing to the effect (see below for further details).

It is possible that with better understanding of the neurological and electrophysiological correlates of attention, that methods may be developed to assess attentional biases using neuroimaging data or event-related potentials (ERPs). ERP studies measure the electrical activity on the scalp caused by a small area of the cortex following the presentation of a stimulus. Characteristic patterns of positive and negative changes in voltage above and below a resting state indicate processing, and a positive wave occurring approximately 300 ms after stimulus presentation has been found to be associated with attention, and in particular the detection of salient events. Studies have shown associations between craving and the size and timing of these components of ERPs to substance-related pictures for heroin users, smokers, and alcoholics.

Models of Attentional Bias

In the use of the visual probe task in anxiety, it has been established that attentional biases can be due to a combination of pre-attentive, attentional, and post-attentional factors.

Pre-attentive effects occur very rapidly and can be detected with subliminal cues, typically presented for less than 50 ms. These effects serve to prepare a shift of attention toward the location of the relevant cue, and so increase the speed with which attention can subsequently be directed toward a probe occurring in the same location as the salient cue (in the attention literature, moving attention toward an object or event is known as engagement, so these pre-attentive effects occur before engagement but facilitate it). Anxiety produces reliable, if small, pre-attentive effects, but it seems that general motivational states such as hunger do not. This could be due to a lower salience of food-related stimuli for hungry participants, compared to negative or threatening words for anxious participants, or it could reflect an underlying difference in the cause of attentional biases in addiction and anxiety.

Attentive effects include engagement, or covert shifts of visual attention toward the salient cue location before the probe appears, as well as overt shifts in gaze direction, and can be found with brief supraliminal cues, typically presented for 50–500 ms. These effects seem to occur for a wide range of cues, including food-related

words and pictures for people deprived of food, body shape, and weight cues for people with eating disorders, and substances related cues for users of those substances.

Post-attentional effects, found in studies with longer cue durations, counteract initial shifts of attention toward a stimulus that people are trying to avoid, or disengagement. An increase in disengagement leads to a bias away from the salient stimulus location, which coupled with subsequent inhibition of return effects, leads to a slowing in detection of a collocated probe (inhibition of return is an attentional effect that makes it hard to attend to a stimulus or location that one has already engaged with and disengaged from). Strictly speaking, disengagement refers to the attentional effects rather than gaze, and may be followed by a shift in covert attention to another location, where it may be engaged on another stimulus.

The relative contribution of these phases of attention to attentional biases was summarized by Weierich, Treat, & Hollingworth as two hypotheses: Vigilance-Avoidance and Attentional Maintenance. Vigilance-Avoidance argues that biases are due to enhanced engagement with a stimulus (vigilance) coupled with delayed avoidance: so substance-dependent or craving individuals would be quicker at allocating their attention toward substance-related cues, and then less able to disengage from them. Attentional Maintenance argues that once salient stimuli have gained attention, they hold it, making it difficult for substance-dependent people to disengage from them. The two ideas are not incompatible, and make similar claims for differences in the pattern of attentional allocation, but deal with different timescales: the overt avoidance of the vigilance-avoidance hypothesis occurs later than the covert disengagement of the attentional maintenance hypothesis. Both hypotheses indicate that people who are actively trying not to attend toward substance-related stimuli, or to look away if they find themselves attending to them, will have relative difficulty employing such a strategy.

Another distinction that can be made takes into account the relative salience of the stimuli. While most anxiety studies have used stimuli that either are threatening, or are not threatening, Wilson and MacLeod used faces varying through five levels from very low to very high threat. When these faces were paired with a neutral face in a visual probe task, they found that while all participants showed attentional bias away from the low threat and toward the very high threat, anxious participants attended to the moderate threat faces, while normal participants avoided them. This, they concluded, supported models in which attentional bias occurs after the salience of the stimulus is assessed, and affects the degree of engagement appropriate to the stimulus. For anxious people, low threats can be avoided but moderate threats become worthy of attention, whereas for normal people, only high threats need to be attended to. To date, varying the salience level of substance-related cues in this way has not been attempted, but it would be a useful way of testing incentive sensitization theory, since this would predict that as individuals become more dependent, the salience of all substance-related cues would change, rather than just the degree of response to cues of intermediate salience.

In their review, Weierich, Treat, and Hollingworth point out that because studies tend to use one (or at most two) cue durations they do not allow any interpretation of the time course of attentional allocation, and so make it difficult to contrast alternative hypotheses about engagement and disengagement, and the role of covert shifts in attention as opposed to overt changes in gaze direction. One obvious way of differentiating between overt changes in gaze location and covert changes in attentional engagement and disengagement would be to use eye-tracking techniques. Mogg, Bradley, Field, and De Houwer, for example, measured gaze location during a visual probe task in which smokers were presented with smoking and neutral pictures as cues for 2000 ms, in a task where they were not instructed to maintain gaze upon the central fixation cross. They found that smokers made their initial shift of gaze away from the fixation cross more often (54% of the time) toward the smoking-related picture, and then looked at the smoking-related pictures for longer, while non-smokers made initial shifts equally to smoking and neutral pictures, and looked at both for the same length of time. Unsurprisingly, the smokers in this study also responded faster to probes replacing the smoking pictures than the neutral pictures, whereas the non-smokers did not. The attentional bias measure in this study might simply reflect the fact that participants who are looking at a picture that is then replaced by a dot probe find it easier to detect and respond to that probe. An alternative approach would be to exclude trials on which participants made eye movements, and to see if there is still a location effect. If a difference can still be found, this would mean that overt attentional shifts are not necessary to facilitate probe detection, and that covert attentional biases are responsible.

Attentional Retraining

In an attempt to resolve the controversy about the causal role that attentional biases play in anxiety, MacLeod and colleagues used a visual probe task to induce an attentional bias toward threat cues. Half of their participants completed 576 trials of a normal

visual probe task, in which the dot that they had to detect followed threat cues and neutral cues with equal probability. The other half of the students completed 576 attentional retraining trials in which the dot probe always followed a threat word. This manipulation successfully trained the experimental participants to allocate attention toward the threat cue: it created an attentional bias toward threat. The manipulation did not affect participants' mood, but when they were then subjected to a mildly stressful task, the experimental group responded with higher levels of anxiety.

In a subsequent study, the contingencies during the attentional retraining trials were reversed to induce an attentional bias away from the threat cues, with correspondingly opposite results: participants now showed a less anxious response to a stressful event. To show the potential therapeutic role that such manipulations might have, this study was carried out using a real world stressor: the participants were emigrating from Singapore to Australia to enroll at University, and they completed the training online in their own homes once a day for 15 days prior to leaving (*see* Technology-Delivered Treatments for Substance Use Disorders: Current Status and Future Directions by Bickel for accounts of computer-based therapies).

Other researchers have used the attentional retraining techniques to modify social anxiety and general anxiety. Along with techniques that address biases in memory encoding, retrieval, and inference, attentional retraining is grouped under the general term Cognitive Bias Modification and now forms a component of cognitive behavioral therapies (CBT; *see* Cognitive Behavioral Therapies by Carroll). Attempts to apply the procedure to retrain attentional biases in substance abusers have reported limited success, however. Field and Eastwood successfully induced attentional biases toward and away from alcohol related cues, and the attentional bias group reported a greater urge to drink and also consumed more in a post-induction test. However, the avoid group did not show a reduction in urge to drink, and neither group showed changes in a general craving measure. Other studies confirmed the power of the visual probe task to induce an attentional bias toward or away from substance-related stimuli, but limited or weak generalization of the bias to stimuli other than those used in the training set, and no generalization to other tasks.

Subsequent studies have increased the number of stimuli, have used repeated training sessions, have been coupled with motivational enhancement, and have had more success. However, there is some doubt about the role that attentional biases alone are playing in the modification of craving and behavior, since the increased intensity of the retraining is very apparent to participants, and induces a great deal of additional processing and awareness. Indeed, emphasis has now moved on to other cognitive biases, in the form of Approach-Avoidance tasks, in which participants have to make overt physical movements toward or away from stimuli (this work is described in Implicit and Associative Processes in Addiction by Wiers & Stacy). While these tasks share some conceptual similarity with attentional bias tasks, they are clearly different and involve many other aspects of cognition, including conscious strategies for controlling and overriding automatic behavior.

Overall, the evidence to date gives attentional biases a real but limited role in addiction. Craving, and hence motivated substance use, can be brought about by an increased sensitivity to substance-related cues, but involves more explicit controlled processes that may themselves affect responses on the tasks conventionally used to measure attentional bias. Whereas attentional retraining tasks do have an effect upon craving and substance use, they may work more through their effect upon later strategic responses to attended cues than upon early and automatic attentional biases.

List of Abbreviations

CBT	cognitive behavioral therapies
ERP	event-related potential
ITI	inter-trial interval

Glossary

Addiction in this chapter, addiction is generally given a wider definition than the clinical diagnoses of Substance Abuse, Dependence, or Harmful Use. For the purposes of experimental cognitive research, 'addicted' individuals may only show some features such as mild tolerance or withdrawal, or may self-identify as addicted or problematic users.

Addiction Stroop a measure of attentional bias in which participants name the color in which substance-related and neutral words are printed, with slower color naming indicating a difficulty in ignoring the meaning of the words (see Fig. 45.1).

Attentional retraining modification of attentional biases through repeated sessions of tasks that require substance users to ignore substance-related cues.

Elaboration a cognitive response to intrusive substance-related thoughts that involves controlled access to memory resources and the creation of sensory mental imagery of substance acquisition and use.

Subliminal and supraliminal a stimulus in the environment or a cognitive experiment is defined as supraliminal if its content can be reported (i.e. it is above a subjective threshold). Subliminal stimuli, in contrast, may be too briefly or faintly presented to be reportable (they are below the subjective threshold), but might still lead to detectable behavioral differences (i.e. they can be above the objective threshold).

Visual probe task a measure of attentional bias in which responses to probes are faster if they are presented in the same spatial location as previous salient cues rather than a simultaneous neutral cue (see Fig. 45.2).

Further Reading

Cox, W.M., Fadardi, J.S., Pothos, E.M., 2006. The addiction-stroop test: theoretical considerations and procedural recommendations. Psychological Bulletin 132 (3), 443–476.

Field, M., Cox, W.M., 2008. Attentional bias in addictive behaviors: a review of its development, causes, and consequences. Drug and Alcohol Dependence 97, 1–20.

Field, M., Munafò, M.R., Franken, I.H.A., 2009. A meta-analytic investigation of the relationship between attentional bias and subjective craving in substance abuse. Psychological Bulletin 135 (4), 589–607.

Kavanagh, D.J., Andrade, J., May, J., 2005. Imaginary relish and exquisite torture: the elaborated intrusion theory of desire. Psychological Review 112, 446–467.

MacLeod, C., Koster, E.H.W., Fox, E., 2009. Whither cognitive bias modification research? Commentary on the special section articles. Journal of Abnormal Psychology 118 (1), 89–99.

Weierich, M.R., Treat, T.A., Hollingworth, A., 2007. Theories and measurement of visual attentional processing in anxiety. Cognition & Emotion 22 (6), 985–1018.

CHAPTER

46

Sensory Imagery in Craving

Jackie Andrade

University of Plymouth, Drake Circus, Plymouth, UK

OUTLINE

WHAT IS SENSORY IMAGERY?

Advertisers have long known the power of mental imagery in creating and sustaining desires. Phrases such as "sun-kissed beaches" and "turquoise waters" help us to conjure up an image of an island paradise and stimulate our desire for a relaxing holiday. We mentally enjoy the beautiful view, feel the warm sand between our toes, and hear the gentle splash of waves on the shore. "Sensory imagery" refers to this ability to mentally simulate the sight, sound, smell, taste, and feel of an experience.

Sensory imagery is used in clinical settings and in laboratory research to stimulate or exacerbate craving for drugs. For example, smokers crave cigarettes more after being asked to imagine relaxing with friends who are lighting up and to imagine the taste and effects of smoking, than after imagining a neutral scenario unrelated to smoking. More vivid imagery of the smoking scenarios is associated with stronger urges to smoke. Likewise, alcohol craving can be induced by asking clients to imagine entering their favorite bar, ordering, holding and tasting a cold, refreshing glass of their favorite beer.

Sensory imagery is a powerful tool for inducing craving because it taps into the cognitive system that underpins human motivation and supports complex behaviors. Imagery plays an important role in motivation because it conveys the emotional qualities of the desired event, mimicking anticipated pleasure or relief.

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00046-2

THE ROLE OF SENSORY IMAGERY IN MOTIVATION

Imagery and Emotion

Sensory images are emotive; imagining a happy event makes us feel happier, and imagining a stressful event makes us anxious. Imagining events has greater effects on emotion than merely thinking about them. Vivid, intrusive imagery is a symptom in a wide range of psychological disorders, from posttraumatic stress disorder to depression and bipolar disorder, obsessive-compulsive disorder, and a variety of anxiety disorders, and is thought to contribute to the maintenance of those disorders. A new line of research and therapy in clinical psychology is showing how psychological disorders can be remedied by "imagery rescripting," encouraging clients to modify their images of experienced traumatic events or feared future events. Laboratory evidence, discussed below, suggests that sensory imagery is also a useful target for tackling craving in addiction, though to date the focus has been on preventing or disrupting imagery.

Elaborated Intrusion Theory

The role of imagery in motivation is explained by elaborated intrusion theory (Fig. 46.1). According to this theory, desires can be triggered by a wide range of stimuli, which includes physiological cues (salivation, stomach rumbles), negative mood, environmental cues, and related thoughts. These triggering stimuli increase the likelihood that we will experience a conscious thought about the target of desire. The thought is intrusive in the sense that it interrupts one's train of thought, but in many cases it may be fleeting, for example, when a momentary thought about lunch interrupts an interesting conversation. Intrusive thoughts become desires when they are elaborated, as might happen when the thought about lunch interrupts one's attention to a dull lecture. Elaboration includes planning ways of achieving the desire ("I could buy a sandwich"), generating expectancies about satisfying the desire ("I'd be able to concentrate better once I've had something to eat"), and thinking about one's self-efficacy or ability to obtain the desired object or activity ("I'm sure I could find a shop that's still open."). Importantly, though, elaboration involves sensory imagery. We imagine engaging in the desired activity and the image conveys some of the pleasure or relief of the real thing. We imagine the sort of food we will buy, the smell and taste of it, and the satisfaction of relieving our hunger. More vivid and realistic images convey greater pleasure but they also make us more acutely aware of the separation between our current state and desired state. Thus desire imagery is briefly pleasurable but ultimately aversive. It is the negative emotion it causes that spurs us to action to achieve our desire and change our current situation.

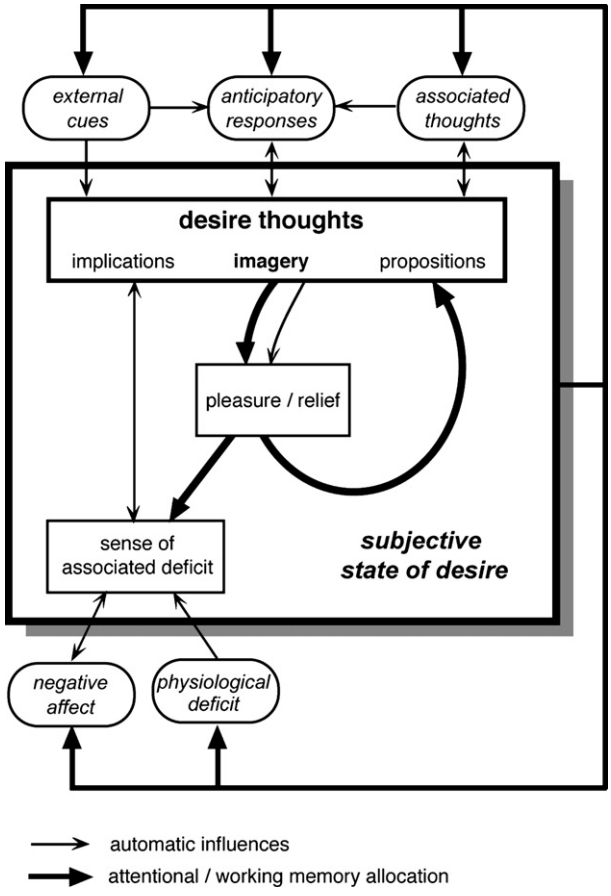


FIGURE 46.1 The elaborated intrusion theory of motivation, showing the contribution of triggers (rounded external boxes), intrusive thoughts ("desire thoughts"), and sensory imagery to craving (central square box). Thick arrows show the controlled processing cycle of conscious imagery and associated affect; thin arrows represent automatic influences on desire. *Adapted from Kavanagh, D. J., Andrade, J. & May, J. (2005). Imaginary relish and exquisite torture: The Elaborated Intrusion theory of desire, Psychological Review, 112(2), 446–467*

SENSORY IMAGERY AND CRAVING

Automatic and Controlled Cognitive Processes in Addiction

Addiction to substances develops through sensitization of the brain's reward system to drug-related stimuli (see Chapter 39 by Robinson, Berridge, & Robinson), and development of highly conditioned responses. Sometimes, exposure to conditioned stimuli is sufficient to trigger drug use with minimal conscious oversight and little or no craving, as when a smoker absent-mindedly lights another cigarette before finishing the current

one. In this situation substance use is automated, with behavior driven by automatically activated action schemas without intervention from top–down control processes.

Such situations are rare, because substance use usually requires more complex sequences of behaviors over extended time periods. Conscious mediation is needed to orchestrate these sequences of behavior and to carry them out in the face of competing goals. So, a typical instance of substance use might involve planning (a telephone call to a drug dealer, a trip to a bar or shop, a break from work), assessing self-efficacy ("could I resist temptation?"), generating expectancies ("a drink would feel great now"), and inhibition of competing goals (to be healthier, to save money, to finish a piece of work). Sensory images are the medium by which we keep in mind the goal of obtaining a drug during this extended sequence of thoughts and actions, and the medium by which we compare competing goals.

Elaborated Intrusion Theory Applied to Addiction

Elaborated intrusion theory views sensory imagery as a key component of consciously mediated substance use behaviors. As with other desires, sensory imagery serves as a mental bridge between our initial thoughts about drug use and actual consumption, and allows us to mentally compare conflicting goals. Craving is the cognitive-emotional state that pertains when sensory imagery of drug consumption is accompanied by felt relief or pleasure and also aggravated feelings of deficit. We shall consider three phases in craving:

Triggers

Craving for drugs can be triggered by drug-related stimuli in the environment (e.g. drug paraphernalia), related thoughts (e.g. thinking about a friend with whom one uses drugs), physiological withdrawal symptoms, anticipatory responses such as salivation, and negative mood. Negative mood is strongly associated with craving and may become increasingly important as addiction develops and negative mood is increasingly attributed to lack of the drug. Development of attentional biases in addiction increases the likelihood that drug-related stimuli will trigger the intrusive thoughts that lead to substance craving (see Chapter 45 by May).

Intrusive Thoughts

Craving triggers do not generally lead directly to consumption but instead increase activation of drug-related information in memory, which in turn increases the likelihood of experiencing a thought about drug use. These intrusive thoughts may be verbal thoughts or image fragments, and may be fleeting. In a study of clients beginning clinical treatment for alcohol dependence, we found that the majority of the sample occasionally experienced thoughts about drinking that popped into mind and then vanished spontaneously. Trying to suppress intrusive thoughts about substance use tends to exacerbate rather than reduce their occurrence.

Elaboration

Thoughts about drug use trigger craving when they are elaborated. In this sense, intrusive thoughts are the gateway to craving. As described above, elaboration includes planning acquisition of the drug, evaluating self-efficacy in resisting craving, comparing conflicting goals, and generating expectancies of the effects of consumption. The key component though, and the one that gives elaboration of its emotional bite, is sensory imagery.

Imagining the sight, smell, taste, and bodily effects of consuming the drug mimics the experience of actually consuming it. This imagery supports the other aspects of elaboration and provides a mental bridge between the trigger and eventual behavior. Imagery sustains motivation in this way because it is emotionally charged. When a smoker imagines smoking a cigarette, they are not merely imagining what the cigarette would look like and taste like, they are also imagining how pleasurable the taste and effects of inhalation would be. An alcoholic anticipating the first drink of the day not only visualizes the drink but, by doing so, also experiences the same pleasure or relief expected from actually drinking it. The pleasure or relief felt when imagining substance use may be weaker than the real thing – a mere simulacrum – but at times it may be even stronger because imagery allows us to mentally conjure up our ideal experience, where substance use is perfectly rewarding and lacks the inconveniences of everyday life, such as having to stand outside in the rain to smoke a cigarette.

The pleasure or relief embodied by the initial image encourages further imagery, but imagery also has aversive effects. Imagery enhances awareness of deficit, as when a smoker imagines the satisfaction of taking a drag on a cigarette and, by comparison of their imagined and actual state, feels more keenly their current discomfort and lack of satisfaction. The ensuing worsening of mood stimulates further, increasingly vivid and realistic imagery in a downward spiral of briefly pleasurable imagery leading to enhanced awareness of deficit and worsening mood.

Competing Goals

Substance use often persists in the face of good intentions to quit. The central role of sensory imagery in craving and substance use helps to explain the relative

weakness of our good intentions. Goals to change our lives, for instance to become healthier, can usually be imagined less vividly than can habitual behaviors, because we have less information in memory to draw on, and in particular little sensory information on which to base an image of an ideal future self that we have never before experienced. This lack of sensory information means that goals to cope with withdrawal symptoms, to resist temptation, or generally to become healthier are difficult to imagine vividly. Images of habitual substance use are more vivid and therefore feel more salient or urgent when weighing up whether to give into an intrusive thought about substance use or to resist temptation for better future health.

Experiences of Sensory Imagery and Craving

Sensory imagery is a consistent feature of people's reports of their craving experiences. Imagery (of sights, tastes, smells, physical sensations, and sometimes sounds) is reported as a feature of craving for a range of addictive substances including alcohol and nicotine, and nonaddictive substances and activities such as snacks and sport. Elaborating sensory imagery not only extends episodes of craving, it makes the craving more intense as the images become more articulated and vivid. Particularly in the case of visual imagery, the strength or vividness of imagery, and the frequency of imagery, correlates with craving strength. For example, most people in our clinical sample with alcohol dependence problems had sensory imagery during their strongest craving episodes when they were trying to control their alcohol use, with their imagery combining more than two senses on average. The most common elements of alcohol imagery were imagining tasting a drink, picturing the drink, and imagining swallowing alcohol. As predicted by elaborated intrusion theory, frequent sensory imagery was associated with stronger and longer-lasting episodes of craving for alcohol.

The cognitive psychology of mental imagery explains how imagery can be prevented or weakened, while research into craving shows how preventing or weakening mental imagery can lead to reductions in craving.

COGNITIVE PSYCHOLOGY OF IMAGERY – CLUES TO TACKLING CRAVING IMAGERY

Mental imagery involves activation of the same brain regions that are activated by actual perception, thus neuroimaging studies show increased activity in visual areas of the brain when participants are seeing an object and when they are picturing an object. This mental "seeing" also requires retrieval of information from long-term memory (generic knowledge of an object's shape and color, for instance, and specific memories of having seen such objects before), generation of an image from the information thus retrieved, and maintenance of that image in consciousness. Retrieval, image generation, and maintenance are carried out by a set of cognitive processes that are collectively known as "working memory." Neuroimaging studies show increased activation in brain regions that support working memory during substance craving (see Chapter 47 by Paulus). To date though, brain imaging has not been used specifically to test the prediction that sensory areas underpinning imagery are activated during craving.

Working Memory

Working memory is the means by which we are able to keep information consciously in mind while transforming it or using it to achieve some goal. For example, when mentally adding up a shopping bill we use working memory to hold or "carry" subcomponents of the sum; when planning a route, we use working memory to visualize landmarks and links between them. In craving, working memory is involved when we recollect where we put our cigarettes, work out if we have time to stop and smoke one, and visualize taking a cigarette out of its packet and lighting it.

Subsystems of working memory are specialized for processing auditory or visuospatial information. Imagining the appearance of a bartender pouring a drink uses visuospatial working memory while imagining the clatter of coins as you win the slot machine jackpot uses auditory working memory. There is a limit to how hard these working memory systems can work. Using visual working memory to remember and tap a pattern on a keyboard, for example, makes it harder to imagine a visual scene such as a busy train station or a rose garden, and such images are rated as less vivid than when participants are doing an auditory task that does not impinge on the visual working memory processes needed for visual imagery. This limited capacity is seen across a range of imagery tasks, from imagining verbally cued neutral scenes as in this example, to keeping recently perceived stimuli vividly in mind, to recollecting personal experiences. Even highly emotive autobiographical memories feel less vivid and less distressing (or less pleasurable) if they are recalled while performing a task that loads working memory.

Working Memory and Craving

Working memory is involved in all aspects of craving, from the elaboration of plans to acquire a drug to trying

to focus on competing goals and inhibit thoughts about drug use. Not surprisingly, people perform less well on a range of working memory tasks when they are craving than when they are satiated. There is a two-way relationship, thus people crave food and drugs less strongly when they carry out tasks that require working memory than when they have nothing else to do.

For clinical purposes, maintaining a perpetual, general working memory load sufficient to block craving is impractical, because it would block most normal cognitive activity as well as craving. Instead, this area of craving research has focused on simpler tasks that selectively compete with aspects of working memory needed for craving imagery. The aim is to blunt the craving or provide a brief respite rather than eliminate it altogether.

INTERFERING WITH CRAVING

The limited capacity of working memory is the clue to reducing craving by reducing craving imagery. As the visuospatial and verbal subsystems of working memory can only hold and process a certain amount of information at a time, occupying them with a simple visual or verbal task is sufficient to inhibit visual or auditory imagery (imagery of sounds). In most substance use, auditory imagery is less likely than visual, gustatory, or olfactory imagery, unless sounds are highly associated with the person's preferred substance (e.g. the sound of popping champagne). Experimental studies have often therefore compared the effects on craving of performing visual working memory tasks versus auditory or verbal tasks that control for generic working memory factors and allow us to test the specific effects of blocking visual imagery. Tasks that involve manipulating, remembering, or imagining visual information make it difficult to visualize drug use simultaneously, and have been shown to lead to lower craving than verbal tasks that permit concurrent sensory imagery.

Cigarette Craving

Imagining neutral scenes (e.g. a rose garden) as opposed to neutral sounds (e.g. a telephone ringing) reduces cigarette craving in abstinent smokers to the levels reported by smokers who are allowed to smoke ad lib before the experiment. Imagining neutral odors is also effective. Blockade of craving imagery need not be achieved through competing imagery; it can also be achieved with tasks that require the same visuospatial working memory processes that are needed for imagery. For example, the development of cigarette craving is inhibited by asking smokers to create specific forms out of modeling clay or plasticine, for example, spheres, cubes, or animal shapes. This clay modeling task is assumed to involve maintenance of a visual representation of the intended shape in working memory plus spatiomotor control processes that are also involved in rehearsal, or "replaying," of visuospatial information in working memory.

Food Craving

Effects of selective cognitive blockade on craving have also been tested in relation to craving for food, either with hungry participants craving food in general or with participants craving chocolate, which is associated with rather specific cravings that cannot be satisfied by nutritionally similar foods. This research has shown that craving can be reduced by imagining neutral scenes, by watching a rapidly changing visual display, and by performing tasks that require spatiomotor control as well as visual working memory processes, for example, making side-to-side eye movements or tapping a pattern on a keyboard. The similarity in effects on food craving and cigarette craving supports the argument that substance craving in general can be reduced by selectively blocking visuospatial working memory.

DIRECTIONS FOR TREATING ADDICTION BY TACKLING CRAVING

Craving is an important target for addiction treatments because it helps to sustain addictive behaviors (more cigarettes are smoked "to relieve craving" than for any other reason) and makes quitting difficult. Intense craving can occur long after physiological withdrawal symptoms have faded (e.g. because of exposure to previously conditioned cues), inducing discomfort and distress and increasing the likelihood of relapse. Psychological treatments are effective generally in addiction, both on their own and combined with pharmacological treatments, but there is much room for improvements in treating craving during the quit attempt and providing better tools to help clients cope with recurrences of craving after quitting.

Imagery Interventions

Because craving, as opposed to addiction, is essentially a conscious, cognitive phenomenon, it is susceptible to interference from ongoing cognitive activities. The laboratory studies discussed above suggest a range of simple tasks that might be useful for controlling craving, tasks that selectively engage processes of visuospatial working memory. Extrapolating to everyday life,

craving should be reduced by spending a few minutes vividly recollecting a happy memory or scene, playing a computer game such as Tetris™ or Puzzle Express™ that requires mental manipulation of visual stimuli, making models from a piece of plasticine, or even watching a screensaver with high contrast, continually changing visual properties.

However, strategies that only target craving imagery may be less than optimally effective because they require a choice to be made, on experiencing an intrusive thought about drug use, about whether to take the drug or try to alleviate craving by performing the imagery-blocking task. People may choose to take the drug because the initial thought is particularly appealing or intrusive, or because they do not want to be distracted from the task in hand – stopping work for a cigarette may seem less distracting than stopping to carry out the intervention or trying to continue working in the face of frequent intrusive thoughts about smoking.

More effective strategies are likely to be those that combine imagery blockade with techniques that interrupt the pathway from initial craving trigger to intrusive thought to elaboration. Simply telling oneself not to think about drug use is counterproductive, as monitoring one's success at not thinking actually triggers thoughts about drug use ("Have I thought about smoking recently? Oops, now I have"). Research with nonclinical samples has shown that telling oneself to ignore, rather than suppress, intrusive thoughts is more effective. Clinical approaches that encourage people to ignore or accept intrusive thoughts are effective across a wide range of disorders where intrusive thoughts are problematic and hold promise for treating addiction and craving.

Thought Acceptance and Mindfulness

Acceptance therapies encourage individuals to view thoughts as transient events and to accept that unwanted thoughts will sometimes pop into mind. Accepting that thoughts come and go again unbidden breaks the cycle of attempted thought suppression, increased thoughts, feelings of failure or rumination on the thoughts, followed by renewed but counterproductive efforts at suppression. In terms of addiction, the aim is to instill an acceptance of the fact that there is no need to elaborate a thought about substance use, or to act on it, because it will go again of its own accord.

Mindfulness-based therapies extend this acceptance approach by teaching individuals to become aware of their whole, changing, sensory experience. Training in mindfulness aims to increase awareness of all thoughts and sensations, and by doing so to decrease the salience of, and elaboration of, unwanted thoughts. Mindfulness-based interventions also reduce distress about the craving, by helping the person observe their thoughts without engaging with them or being concerned about what they may mean.

Implementation Intentions

A simple but effective way of increasing the likelihood that someone sticks to their good intentions and resists temptations, is to ask them to make an "if…then" plan or implementation intention that specifies when, where, and how they will achieve their goal. For example, someone trying to quit smoking might make the plan that "if my friends start lighting up at tomorrow's party, then I shall chew a piece of chewing gum instead." Specifying how one will achieve a goal in this detailed, concrete fashion has several beneficial effects:

- it increases the salience and memorability of the important cues ("my friends lighting up") and intended behaviors (chewing gum rather than smoking a cigarette)
- it strengthens the link between the cues and the intended behavior, so that seeing someone else smoking automatically cues thoughts about chewing gum
- it may also encourage sensory imagery of the desired outcome, so that the abstaining smoker has a ready-made image of themselves in a specific situation taking out a piece of gum and chewing it, reducing the likelihood that they start imagining themselves smoking a cigarette instead.

Implementation intentions may thus bypass elaboration of thoughts about drug use, or at least strengthen the vividness of imagery of the intended rather than habitual outcome. Although this aspect of implementation intentions has received little research attention, there is recent evidence that explicitly encouraging participants to imagine carrying out their "if…then" plan enhances the efficacy of implementation intentions.

FUTURE DIRECTIONS FOR CRAVING INTERVENTIONS

Mindfulness training includes strategies such as body scanning that help increase awareness of experiences and thoughts other than the unwanted ones. Body scanning involves shifting attention from one part of one's body to another, a process that is likely to involve visual imagery of the attended parts of the body. Recent studies with nonclinical populations have shown that body

scanning reduces intrusive thoughts about food as effectively as a distraction technique involving guided sensory imagery of a woodland walk and that it reduces smokers' desire to smoke compared with listening to a natural history tape.

Interventions like body scanning that combine positive or neutral sensory imagery, or blockade of visuospatial working memory, with acceptance of intrusive thoughts should be particularly effective at reducing drug craving because they should reduce the likelihood of intrusions being elaborated while simultaneously reducing the vividness of any sensory drug imagery that does occur, breaking the vicious circle whereby intrusive thoughts trigger sensory imagery which in turn triggers further thoughts. Helping clients to create positive images of abstinence behaviors should also be effective in helping to bypass the process whereby initial thoughts about substances lead to sensory imagery of substance use. Future research might trial combinations of thought acceptance and visuospatial working memory tasks, or test the effect of including a vivid, pleasant image in an implementation intention plan.

Reduced sensory imagery of substance use may be playing a role in another intervention that is currently attracting research attention. Brief periods of moderate physical exercise have been shown to be effective at reducing desire to smoke. The mechanisms are not well understood and may include effects on mood and confidence as well as attention to the benefits of health and the debilitating effect of smoking on lung function. General distraction is not the mechanism, but competing visual distraction or visual imagery may be contributing to the effects, particularly in the case of exercise in natural settings rather than the gym.

CONCLUSIONS

Sensory imagery is central to craving, giving craving its emotional bite and prolonging episodes of craving in the face of competing desires to abstain. Stronger, more vivid imagery predicts stronger craving in clinical and nonclinical samples, while blocking imagery with simple working memory tasks has been shown to reduce craving. Studies of food craving and of smokers' craving for cigarettes suggest that techniques combining thought acceptance and competing sensory imagery will be effective for reducing the frequency and strength of episodes of craving. Self-report data show similar craving phenomenology across addicted and nonclinical samples; therefore, we predict similar effects of these techniques in clinical populations. Effective interventions for substance craving in addiction are likely to be those that combine strategies for disrupting the elaboration of intrusive thoughts about substance use with strategies for blocking any ensuing sensory imagery.

SEE ALSO

Cognitive Factors in Addictive Processes, Implicit and Associative Processes in Addiction, Neural Correlates of Craving for Psychoactive Drugs, Relation of Craving and Appetitive Behavior, Attentional Biases in Craving

Glossary

Desire a state in which thoughts and images about a target behavior carry emotional weight, both pleasure or relief associated with achieving the target and negative affect from awareness that it has not yet been achieved. Cravings are desires to consume substances.

Elaborated intrusion theory a theory of human motivation in which desires are sustained by emotionally charged sensory imagery

Elaboration mentally embellishing an initial thought or stimulus, for instance a thought about alcohol might be elaborated by retrieving memories of yesterday's drinking session, imagining the taste of beer, or planning a visit to a bar.

Implementation intention a plan stating how, where, and when one will perform a desired behavior

Intrusive thought a seemingly spontaneous thought that interrupts our stream of consciousness

Sensory image a mental representation or simulation containing information about appearance, sound, smell, taste, or physical sensations

Vividness the extent to which sensory imagery mimics actual experience

Working memory the mental processes by which we temporarily store and manipulate information to perform tasks such as mentally planning a route, taking part in a conversation, or imagining smoking a cigarette

Further Reading

Baddeley, A., Andrade, J., 2000. Working memory and the vividness of imagery. Journal of Experimental Psychology: General 129, 126–145.

Cropley, M., Ussher, M., Charitou, E., 2007. Acute effects of a guided relaxation routine (body scan) on tobacco withdrawal symptoms and cravings in abstinent smokers. Addiction 102, 989–993.

Holmes, E.A., Mathews, A., 2010. Mental imagery in emotion and emotional disorders. Clinical Psychology Review 30 (3), 349–362.

Kavanagh, D.J., Andrade, J., May, J., 2005. Imaginary relish and exquisite torture: the Elaborated Intrusion theory of desire. Psychological Review 112 (2), 446–467.

Kavanagh, D.J., May, J., Andrade, J., 2009. Tests of the elaborated intrusion theory of craving and desire: features of alcohol craving during treatment for an alcohol disorder. British Journal of Clinical Psychology 48, 241–254.

Kemps, E., Tiggemann, M., 2007. Modality-specific imagery reduces cravings for food: an application of the Elaborated Intrusion theory of desire to food craving. Journal of Experimental Psychology: Applied 13 (2), 95–104.

Knäuper, B., Roseman, M., Johnson, P.J., Krantz, L.H., 2009. Using mental imagery to enhance the effectiveness of implementation intentions. Current Psychology 28, 181–186.

May, J., Andrade, J., Kavanagh, D., Panabokke, N., 2010. Visual imagery tasks suppress craving for cigarettes. Behaviour Research and Therapy 48, 476–485.

May, J., Andrade, J., Kavanagh, D., Penfound, L., 2008. Imagery and strength of craving for eating, drinking and playing sport. Cognition and Emotion 22 (4), 633–650.

Murgraff, V., White, D., Phillips, K., 1996. Moderating binge drinking: it is possible to change behaviour if you plan it in advance. Alcohol and Alcoholism 31, 577–582.

Tiffany, S.T., Hakenewerth, D.M., 1991. The production of smoking urges through an imagery manipulation: psychophysiological and verbal manifestations. Addictive Behaviors 16, 389–400.

CHAPTER

47

Neural Correlates of Craving for Psychoactive Drugs

Jennifer L. Stewart, Martin P. Paulus

University of California, San Diego, CA, USA

OUTLINE

INTRODUCTION

Craving consists of two fundamental processes and can be defined as: (a) a conscious, intense emotional feeling that is associated with (b) a strong urge to act to obtain and use drugs. More specifically, over time a stimulus associated with prior drug use triggers brain mechanisms associated with a physical need, or urge, to respond to the stimulus in drug-addicted individuals. This bodily urge is then translated into a targeted goal to obtain and use the drug. Actions are deployed to fulfill the urge and evidence is collected and evaluated

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00047-4

by neural systems to determine the relative outcome (success or failure) of the urge-related action. Finally, the sensory system determines whether the urge was satisfied.

Craving and urges depend upon the current body state of the individual; for example, the strong urge to eat occurs more frequently in a hungry individual or the need to seek shelter from the cold is particularly strong in a person experiencing hypothermia. Thus, craving can be considered a part of a homeostatic regulatory process. Cravings and urges can arise from the positively reinforcing properties of drugs, wherein the individual experiences cravings to feel better, but they can also emerge due to the negatively reinforcing aspects of the lack of the drug, in instances which the individual experiences cravings to avoid feeling bad. In addition, stress and negative affect may result in homeostatic imbalances that can intensify craving, disrupt decision-making processes, and reduce control over urges, thereby increasing the probability of drug use.

Craving is thought to be an important factor in the maintenance of drug addiction and the probability of relapse. However, there is no universally accepted definition of craving, and debate exists in the field regarding the optimum way to measure the craving experience. Although the focus of this debate has been on self-report indices of craving, a growing literature demonstrates that the craving experience also encompasses changes in patterns of physiology and behavior that influence the desire to use drugs. In recent years, research examining brain mechanisms involved in craving has benefited from advances in neuroimaging technology, which have fostered considerable progress in the identification of brain regions involved in drug addiction.

Neuroimaging techniques such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) offer whole-brain measurements of regionally specific changes in brain activity that can be measured during the craving experience. In addition, event-related brain potentials (ERPs), indices of cortical electrical activity, have temporal resolution on the order of milliseconds and can be used alone or in conjunction with other imaging methods to better explain how craving-related processes unfold over time.

The present review of craving has two main objectives. The first objective is to examine the status of neuroimaging research on craving – what we have learned and what we have yet to learn – predominantly within the context of stimulant addiction. Our primary focus will be the study of craving in cocaine addiction, since few studies have examined the neural correlates of self-reported craving responses in methamphetamine and other stimulant drugs that have become increasingly popular, such as dextroamphetamine and methylphenidate, or in other drugs such as cannabis and heroin. However, we will also provide relevant examples, when available, from neuroimaging studies of other drugs (with an emphasis on alcohol and nicotine) in order to demonstrate similarities across substances. The second objective is to explain how the study of brain regions involved in interoception, an intrinsic but largely overlooked component of the craving experience, may contribute to knowledge regarding precipitating factors for drug dependence and relapse and improve treatment outcome for addicted individuals. Before discussing neural systems involved in the implementation of the craving response, however, we will first review paradigms utilized in the neuroimaging literature to elicit craving in drug users and the measurements of craving that are then correlated with brain activation during these paradigms.

CRAVING ELICITATION AND MEASUREMENT

In this section, we explain three paradigms used to elicit craving and provide an overview of self-report measures researchers use to determine the extent to which an individual is craving drugs as a result of these paradigms.

Cue-Reactivity Paradigms

Neuroimaging researchers have successfully elicited heightened levels of craving and urges in drug-dependent individuals using cue-reactivity paradigms, imagery scripts linked to an individual's personalized drug use experience, and drug self-administration schedules, although the latter is less frequently utilized. Cue-reactivity paradigms appear to be the most popular tasks utilized in alcohol, cocaine, and nicotine craving research, and within the context of cocaine addiction these paradigms have often involved films depicting an individual or group of users smoking or snorting cocaine. These films are then contrasted with a neutral film, such as those depicting nature scenes, and in some cases, films associated with appetitive motivation (sex, displays of happiness) or aversive motivation (displays of negative affect), in order to show that self-reported craving is higher in the cocaine film than the other films and then demonstrate that the neural mechanisms of craving are specific to cocaine and not to other types of emotional experiences. In addition to films, pictures of cocaine paraphernalia (e.g. glass pipe, mirror, razor blade, straw, and dollar bill) have been shown in conjunction with audiotapes of individuals simulating the use of cocaine or describing their experience with the drug. One study

allowed cocaine users to touch drug paraphernalia and compared brain activation and craving levels with conditions wherein cocaine pictures or audiotapes were played.

Personalized Imagery Scripts

Similar to cocaine stimuli used in cue-reactivity tasks, personalized imagery scripts used to evoke cocaine craving are compared with other personalized scripts describing neutral events or types of negative affect (anger, stressful situations) to determine specificity of brain activation linked to craving. A recent study paired personalized neutral and cocaine scripts with threat of electric shock to examine whether stress intensified brain regions involved in craving specifically during the cocaine script.

Drug Self-Administration Paradigms

Whereas cue-reactivity and imagery paradigms are common in the cocaine craving literature, few studies have examined craving as a function of drug self-administration, potentially due to the ethical limitations of such a design. Although not a study of cocaine self-administration, two additional studies have examined craving in cocaine abusers after administration of another stimulant (methylphenidate) compared to a placebo. In addition, a paradigm relevant but not identical to drug self-administration tasks has been implemented in several studies of nicotine craving, wherein individuals are tested once during drug abstinence and tested again in a satiated state in order to compare changes in craving at rest or in response to drug cues within the context of cue-reactivity paradigms. Overall, the fundamental approach to examine craving has been to change the individual's homeostatic state toward a need state that elicits feelings to approach and actions to acquire the drug of choice.

Self-Report Measures of Craving

In most imaging studies of cocaine craving, brain activation during the cocaine cue or script condition is contrasted with the remaining conditions and correlated with craving self-reports that are administered before, during, or after the imaging session. Debate exists regarding how best to measure craving, and as a result, there is no gold standard self-report index of the experiential craving response. Some researchers have advocated the collection of moment-by-moment measures of craving instead of administering trait levels of craving, since many studies indicate that craving fluctuates significantly even during a 1-day period. For example, an electronic diary study examining the frequency and duration of craving in 112 cocaine users demonstrated that during the 5 h prior to cocaine use, ratings of craving significantly increased, and in periods of verified cocaine use, craving rates increased during the day and were higher than during periods of cocaine abstinence.

Although, in rare instances, cocaine users are selected as participants on the basis of high levels of craving in the previous week, the most common indices of self-report craving in neuroimaging studies of cocaine addiction are Likert or visual analog scales administered immediately before, during, and/or immediately after exposure to cocaine and comparison cues, asking questions regarding, in the present moment, what degree and/or frequency the cues elicited a craving, urge, want, or need for cocaine; whether the participant could resist cocaine if it was offered; or how much the cues excited or energized the participant. A measure used less frequently in imaging studies is the cocaine craving questionnaire (CCQ), a 45-item scale indexing several factors of drug use, such as desire to use cocaine, intention to use cocaine, lack of control over cocaine use, anticipating positive outcomes from cocaine use, and anticipating relief from withdrawal symptoms.

A comparison of cocaine craving measures endorsed by patients in a substance treatment program for cocaine abuse and dependence determined that one-item visual analog and Likert scales are best at predicting positive urine screens and short-term cocaine use, whereas the CCQ is the best predictor of future long-term cocaine use. Since self-reported measures of craving do not always correlate highly with one another nor do they consistently predict drug use and relapse, craving studies in general have been critiqued for not recording multiple measures of the craving response, such as indices of urge-related behavior (how much the individual would be willing to pay for the desired drug) and additional physiological indices of arousal and intensity (skin conductance and heart rate) or affect (electromyographic activity, coding of facial expressions, and startle responses).

Determining the Brain Regions Linked to Indices of Craving

The fundamental goal of neuroimaging research in craving is to relate the intensity of the craving experience to the degree and timing of activation in brain regions thought to be involved in the implementation of craving within the context of addiction. Three possible avenues exist to accomplish this goal. First, verbal behavior associated with the craving experience (such as the desire to use a drug and perceived control

over drug use) can be measured with self-report rating scales. Second, behavioral consequences associated with increased craving (such as stronger motivated actions to acquire drug-related stimuli) can be assessed. Third, nonverbal hedonic changes in the individual (such as facial expression) can be indexed. It is advantageous to use multiple methods to study the neural mechanisms of craving in order to limit the influence of demand characteristics and provide convergent validity for craving as a multidimensional construct involving experiential, behavioral, and physiological changes within the individual.

SPATIAL DYNAMICS OF NEURAL SYSTEMS INVOLVED IN DRUG CRAVING

It has been argued that initial drug exposure activates the mesolimbic dopamine system in subcortical neural systems involved in reward processing, wherein ventral tegmental neurons release dopamine in their projections to the ventral striatum and amygdala. Dopamine release signals other brain regions to develop behavioral action plans in response to the drug, and as a result, learned associations are formed that reinforce the rewarding qualities of drug use. Repeated drug use renders these regions hypersensitive, increasing the drug's incentive salience, or its desirability and ability to command attention and induce appetitive motivation. Moreover, this heightened incentive salience is critical for the expression of urges and craving. Finally, the repetitive emergence of rigid behavioral patterns expressed as habitual, compulsive drug seeking is thought to be mediated by the dorsal striatum. The reinforcing properties of the drug are exaggerated relative to natural reinforcers in the environment. Whereas the influence of the subcortical reward system intensifies during drug use, the contribution of the executive function system of frontal brain regions weakens, which is important for decision making and inhibitory control by signaling future outcomes or consequences (reward or punishment) of drug use. As a result, executive control over drug seeking is diminished and areas of the prefrontal cortex become hyper-responsive to cues predicting drug availability. Figure 47.1 illustrates subcortical and cortical brain regions implicated in the drug craving experience.

Ventral Striatum, Amygdala, and Hippocampus

The ventral striatum, amygdala, and hippocampus are components of the mesolimbic dopamine pathway, which has been implicated in the rewarding effects

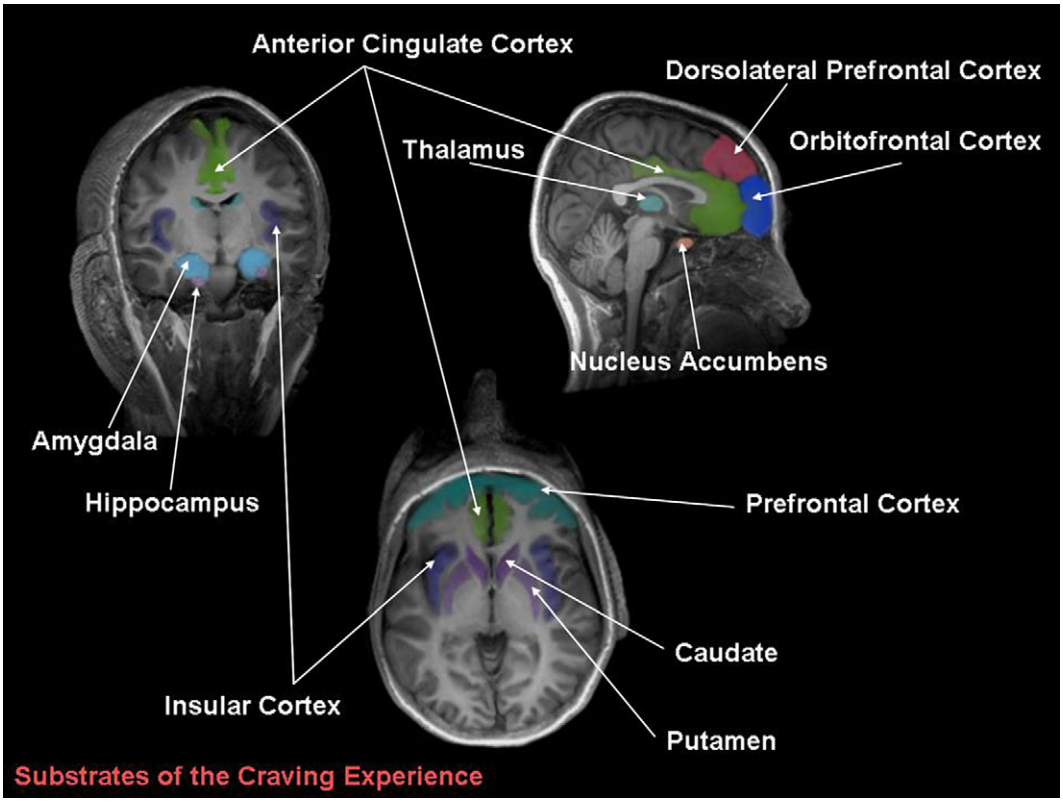


FIGURE 47.1 A primer on the subcortical and cortical substrates of the craving experience.

of drugs, drug-related memories, and conditioned responses. The ventral striatum consists of the nucleus accumbens, which receives dopamine input from the ventral tegmental area as well as input from the amygdala, hippocampus, and prefrontal cortices. Although the ventral striatum is thought to be involved in the rush or high experienced after ingesting a drug, a few studies also implicate this region in the craving process. Two studies of cocaine self-administration have demonstrated that heightened ventral striatum activation was associated with self-reported craving levels in cocaine-dependent individuals, and an additional study found that right nucleus accumbens activation was associated with higher craving ratings during script-guided imagery of acts and sensations linked to cocaine use in a small sample of cocaine-dependent men.

Furthermore, the link between craving and ventral striatum activation has been shown in the nicotine fMRI literature, wherein nicotine-abstinent female smokers exhibited greater nucleus accumbens activation in response to nicotine-related pictures that was associated with baseline pre-fMRI scan craving ratings. Similarly, a recent PET study of alcohol craving indicated that subjective craving ratings in recently detoxified alcohol-dependent men were linked to ventral striatum activation that increased in response to alcohol cues but decreased during the expectation of reward. These alcohol, cocaine, and nicotine data are in line with additional research suggesting that the ventral striatum is involved in the initial rewarding effects of a drug, and neurochemical changes in this region spread progressively to more dorsal regions of the striatum to instantiate a pattern of habitual drug-taking behavior.

In addition to the ventral striatum, the amygdala may contribute to the craving experience by establishing associations between the rewarding effects of drug intake and initially neutral predictors of impending drug administration, such as paraphernalia and individuals associated with drug use, which eventually signal the potential for reward over time. In addition to facilitating conditioned learning, the amygdala may also enhance emotional memory for episodes of and contexts involving drug use. For example, one PET study demonstrated that, compared to a neutral imagery condition, craving ratings during personalized autobiographical memories and sensations associated with cocaine use were linked to greater bilateral amygdala activation (with right amygdala activation larger than the left) in abstinent cocaine-dependent men. Moreover, two additional PET studies have shown that heightened amygdala activation was linked to higher levels of craving in response to cocaine cues (scripts, videotapes, and paraphernalia) in samples of cocaine abusers. The available literature implicates the amygdala in cue and context learning that may facilitate craving processes.

Like the amygdala, the hippocampus may be critical for the formation of associations between drug use and particular spatiotemporal events, which can subsequently serve as stimuli triggering craving. This hypothesis is consistent with alcohol, cocaine, and nicotine research on the craving experience. For instance, an fMRI study examining nonabstinent male smokers with varying levels of nicotine dependence observed a relationship between self-reported craving intensity in response to nicotine-related pictures and amygdala, hippocampus, and parahippocampal gyrus activation. In addition, a PET study indicated that recently detoxified alcohol-dependent men exhibited heightened amygdala and hippocampus activation in response to ethanol odors that correlated with higher craving scores. Although this pattern of results was seen prior to alcohol-related treatment, these correlations did not emerge posttreatment, suggesting that neural indices of craving may be able to predict future use and relapse. Furthermore, an fMRI study that measured brain activation before and after cocaine infusion revealed that higher craving levels were linked to increased bilateral parahippocampal and right hippocampal activation in a sample of cocaine-dependent individuals. In addition to being correlated with craving, right hippocampal and left parahippocampal activations were linked to self-reported "rush" after cocaine administration, suggesting that these regions are involved in the formation of new memories of declarative drug-taking experiences as well as the retrieval of these memories, which may be an important aspect of the craving process.

Dorsal Striatum

Although imaging studies of cocaine addiction implicate both the dorsal and ventral striatum in the craving process, the former has been reported more frequently, particularly within the context of heightened dopamine concentrations in this area. The dorsal striatum consists of the caudate nucleus and putamen and is thought to be involved in the implementation of automatized, or well-learned, habitual drug use urges, cravings, and actions that develop with repeated drug usage. With respect to craving processes, imaging studies involving drug administration have demonstrated a link between dorsal striatum activation and the craving experience that may be associated with habitual patterns of drug use.

For instance, PET research has demonstrated a link between right dorsal striatum activation and cocaine craving approximately 30 min following administration of another stimulant (methylphenidate) in cocaine-abusing individuals. Moreover, a follow-up PET study

demonstrated that during administration of methylphenidate to cocaine users, dopamine levels increased in the dorsal striatum. This relationship held during films depicting cocaine use but not during videos depicting nature scenes, suggesting that dopamine is rendering the dorsal striatum hypersensitive only within the context of drug-related stimuli. In addition to imaging studies of methylphenidate administration, an fMRI study of cocaine self-administration reported a correlation between putamen activation and craving in cocaine-dependent males.

Imaging studies employing drug cue stimuli have also shown that the dorsal striatum is implicated in craving processes. For example, heightened dopamine receptor occupancy in the putamen as measured by PET has been correlated with higher craving intensity during a videotape of cocaine use as well as an audiotape describing pleasurable experiences associated with cocaine use in a sample of frequent cocaine users. Similarly, in response to films depicting users smoking cocaine, cocaine-dependent individuals exhibited increased extracellular dopamine in the caudate and putamen in a recent PET study. This dopamine increase correlated with self-reported craving and was strongest in participants who reported the greatest levels of withdrawal and overall drug addiction severity, suggesting that craving processes become automatized as a function of chronic drug use.

Taken together, findings from this literature support the idea that repeated drug use is linked to dorsal striatum hyperactivity in response to cocaine cues and/or intake, thus increasing salience for the drug that is expressed as craving. However, drug cues or intake may not be required for activation of habitual responses associated with drug use such as craving. For example, an fMRI study demonstrated that right caudate activation was correlated with cocaine craving scores during stress-related imagery in recently abstinent cocaine-dependent individuals, evidence suggesting that the experience of negative affect may be sufficient (and cocaine-related stimuli are not necessary) to trigger regions involved in cocaine habit formation. Although the dorsal striatum – subjective craving link is prominent in cocaine studies, this relationship has not yet been strongly established for drugs such as cannabis, nicotine, or opiates. Additional research is needed to examine this issue.

Thalamus

The thalamus has been viewed traditionally as an important subcortical structure involved in filtering and gating sensory information processing. In the context of drug addiction, thalamic processing may be critical for the selective processing of stimuli that have become associated with drug cues. Thalamus dysfunction may contribute to the craving process by inducing an experience of sensory overload, in which the addicted individual is flooded with stimuli signaling the rewarding properties of drug use, thereby eliciting craving responses. An fMRI study examining differences in craving levels and executive functioning in two groups of cocaine-dependent participants demonstrated that an early-abstinent group (individuals who had been cocaine-abstinent for about a week) exhibited higher thalamus activation and higher craving ratings than a late-abstinent group (individuals who had been cocaine-abstinent for more than 2 weeks) in response to errors during a stop signal task, and this thalamus activation also correlated with craving ratings, suggesting that excessive thalamic responses to conditioned cues may precipitate craving. An fMRI study also showed that higher right thalamus activation was linked to self-reported levels of craving during stress-related imagery in cocaine-dependent individuals who were substance-free for at least 2 weeks. However, studies of cue-elicited cocaine craving have not described a link between thalamus activation and the craving experience, so additional research is needed to clarify the specific role of the thalamus within the context of craving responses.

Dorsolateral Prefrontal Cortex

Whereas heightened activations in subcortical regions such as the thalamus, amygdala, and striatum are implicated in conditioned reward responses that serve to perpetuate drug use, heightened activation in frontal cortical regions such as the dorsolateral prefrontal cortex (DLPFC) during craving may instead reflect a thwarting of cognitive control processes, wherein resources typically used to inhibit behavior resulting in negative consequences are instead channeled into seeking and using drugs despite these consequences.

Although numerous imaging studies of cognition and emotion have linked the DLPFC to executive functioning processes such as planning for future events and maintaining current goals in working memory, much less research exists on the role of DLPFC in the experience of cocaine craving. However, the few available studies on this topic suggest that DLPFC may be lateralized to reflect different reinforcing mechanisms underlying craving, wherein left DLPFC activation is linked with positive drug reinforcement (reflecting the rewarding, appetitive effects of cocaine) and right DLPFC activation is associated with negative drug reinforcement (reflecting avoidance of withdrawal and distress due to homeostatic imbalances and/or lack of recent drug use).

Two PET studies used identical stimuli (cocaine films, scripts, and paraphernalia compared to art-related cues) to examine craving in cocaine abusers, and one of these studies attempted to increase the likelihood of craving by pairing cocaine cue-exposure with the anticipation of drug use (the ability to use cocaine after cocaine cues were presented). Although both studies demonstrated that right DLPFC activation was linked to craving in response to cocaine cues, the study allowing cocaine self-administration showed that left DLPFC activation was also linked to craving responses, suggesting that left DLPFC is recruited to maintain goal set (using drugs) during craving experiences that are linked to an expected reward outcome (having the definitive opportunity to use drugs). Comparable results have been found in an fMRI study of smokers, wherein left DLPFC was associated with smoking image-induced craving as a function of drug use expectancy: the DLPFC activation increased with craving when subjects expected to smoke immediately after their MRI scan, whereas DLPFC activation decreased with craving when subjects knew that they had to wait 4 more hours to smoke after their MRI scan ended. A PET study of nicotine addiction also found heightened DLPFC activity in heavy smokers that linked to self-reported craving in response to handling a cigarette and watching a nicotine video. Left DLPFC activation may be associated with increased appetitive/incentive motivation to obtain drugs, since growing electroencephalography (EEG) and fMRI literatures implicate left prefrontal regions in approach motivation and behavior.

In addition to PET findings of DLPFC activation in cocaine users, one fMRI study demonstrated that right DLPFC activation was correlated with distress and cocaine craving in response to personalized stressful event scripts in recently abstinent cocaine-dependent individuals. Given that several EEG and hemodynamic imaging studies have implicated right prefrontal regions in the experience of withdrawal-related negative emotions, the neuroimaging findings presented above suggest that right DLPFC activity may be linked to a negatively valenced component of the craving process. This negative affect may not be measured by all craving questionnaires or instruments; however, since the majority of imaging studies of cocaine craving have not found a link between the experience of craving and the right prefrontal activation.

Orbitofrontal Cortex

The orbitofrontal cortex (OFC) is thought to play an important role in drug addiction due to its function in the learning of stimulus-reinforcement situations and the prediction of reward. The OFC activation in response to drug-related stimuli may signal prediction of immediate reward of drug use that likely contributes to the craving process. One PET study has indicated that left lateral OFC activation was linked to craving during presentation of cocaine scripts, films, and paraphernalia in cocaine abusers. Likewise, a study demonstrated that heavy smokers exhibited heightened bilateral OFC activity that was linked to self-reported craving in response to handling a cigarette and watching a nicotine video. However, three additional cocaine studies found that right, not left or bilateral, lateral OFC activation was correlated with craving experienced after self-administration of stimulant drugs. For example, PET studies (one with cocaine abusers, one with cocaine-dependent individuals) demonstrated a link between cocaine craving approximately 30 min following administration of another stimulant (methylphenidate). Similarly, an fMRI study also showed that right lateral OFC activation correlated with craving following self-administration of cocaine in cocaine-dependent individuals. The fact that OFC has been linked to craving ratings within the context of definitive future drug administration suggests that this brain region is likely signaling the reward value for situations in which impending drug use is certain.

ROLE OF INTEROCEPTION IN CRAVING: INSULAR CORTEX AND ANTERIOR CINGULATE CORTEX

Models of Addiction Integrating Interoception with Drug Use

Although much of the research spotlight has focused on the role of dopamine in drug addiction, in recent years models of addiction have begun to highlight the role that internal bodily sensations and feelings may play in the maintenance of compulsive drug use. For example, Bechara and colleagues have proposed a somatic marker model of addiction, wherein drug-addicted individuals make poor decisions (continue to use drugs to obtain immediate reward despite negative future consequences) as a result of dysfunctional physiological feeling states expressed in response to rewarding or punishing stimuli. In addition, Siegel developed a conditioned opponent process model, in which he posited that over the course of drug addiction, the body develops negative homeostatic processes to counteract the positive effects of the drug, wherein craving experiences become negatively valenced conditioned responses to drug-related stimuli. Furthermore, in their model of implicit and explicit drug motivational processes, Baker, Curtin, and colleagues have asserted that interoceptive bodily cues are much stronger than

exteroceptive cues in eliciting craving and triggering drug use within the context of addiction.

Interoceptive Processing and Sensations

Implicit in the craving process is the current state of the individual, which is largely a function of the interoceptive experience, the sense of the physiological condition of the entire body. Interoceptive processing involves the ongoing afferent monitoring and efferent modulatory control of inputs arising from the inside of the body. In the process of drug addiction, interoception reflects the experience of intensity or arousal associated with a homeostatic imbalance of the body, resulting in a conscious feeling of the urge to use drugs. Habitual use of a particular drug (e.g. snorting cocaine) is associated with specific interoceptive sensations that are an integral component of the conscious pleasure derived from the drug ritual itself. Research has shown that cocaine-dependent individuals report interoceptive sensations such as restlessness, butterflies in stomach, warm excitement, jitteriness, bodily tension, and rapid heartbeat in response to real-life scenarios involving drug-related cues (e.g. receiving a paycheck, encountering drug dealers or friends who use cocaine, celebrating special occasions, having a surplus of money available) that were accompanied by craving and eventually resulted in cocaine use.

Neural Network of Brain Regions Involved in Interoception

These types of experiences involving somatosensory and autonomic information are relayed from ascending brain areas such as the reticular formation and thalamus to the insular cortex, where they are synthesized and represented as self-aware feeling states. The insula then generates a body prediction error, which is a signal indexing the difference between the individual's current versus predicted bodily state. Next, the insula relays this body prediction error involving emotional feeling states and sensations to the anterior cingulate cortex (ACC). The ACC, in turn, generates an error signal between the correct (or probable) versus incorrect (or improbable) potential outcome, otherwise known as a value signal, to indicate the need to employ attentional resources to adjust a behavior or cognition and thus regain a homeostatic balance (Fig. 47.2). A strongly felt body state (represented by the insula) linked with the incentive to act (represented by the ACC) is an important characteristic of craving. Interoception and its amplification or attenuation by top-down or bottom-up modulation are important determinants of the craving experience.

The insula and ACC both provide information about an individual's interoceptive state to frontal circuitry involved in executive function (DLPFC, OFC) and subcortical circuitry involved in reward processing (striatum, amygdala). The insula possesses reciprocal connections with these regions and is thus centrally placed to receive information about appetitive and aversive salience and the relative value of the stimulus environment in order to integrate this information with the effect that these stimuli may have on the internal body state. Monitoring these bodily sensations is important for the integrity of internal body state and for connecting to systems involved in allocating attention, evaluating context, and planning actions.

Functions of the Insula and Anterior Cingulate Cortex

One function that has been proposed for the insular cortex is to monitor and predict the current versus predicted interoceptive state of the individual, and the role of the ACC is to receive this prediction and generate an "error" or "conflict" signal to initiate new motor behaviors and cognitive plans in order to correct an interoceptive imbalance. Given that the insula registers the homeostatic imbalance of the body – the violation of interoceptive expectancy – it is predicted that higher levels of craving should be associated with greater insula activation in individuals who abuse or are dependent on drugs, reflective of heightened somatic conflict linked to hyperarousal. Similarly, greater ACC activation should also be linked to higher craving in drug users, reflective of heightened attentional resources required to determine action plans in the midst of potentially conflicting alternatives.

NEUROIMAGING EVIDENCE LINKING INTEROCEPTION AND CRAVING

Cue-Reactivity Paradigms

Hemodynamic imaging studies utilizing cue-reactivity paradigms demonstrate a link between cocaine craving and brain regions implicated in interoceptive processing. For example, prior fMRI work has indicated that crack cocaine users exhibited greater left ACC and right insula activation than comparison subjects without drug use in response to a crack cocaine film compared to a nature film. In addition, ACC activation was larger for the cocaine film than a sex film in cocaine users, suggesting that ACC is a brain region specific to craving, not a region simply generalized to reward processing. Furthermore, PET research comparing cocaine cues to art cues using audio scripts, videotapes, and

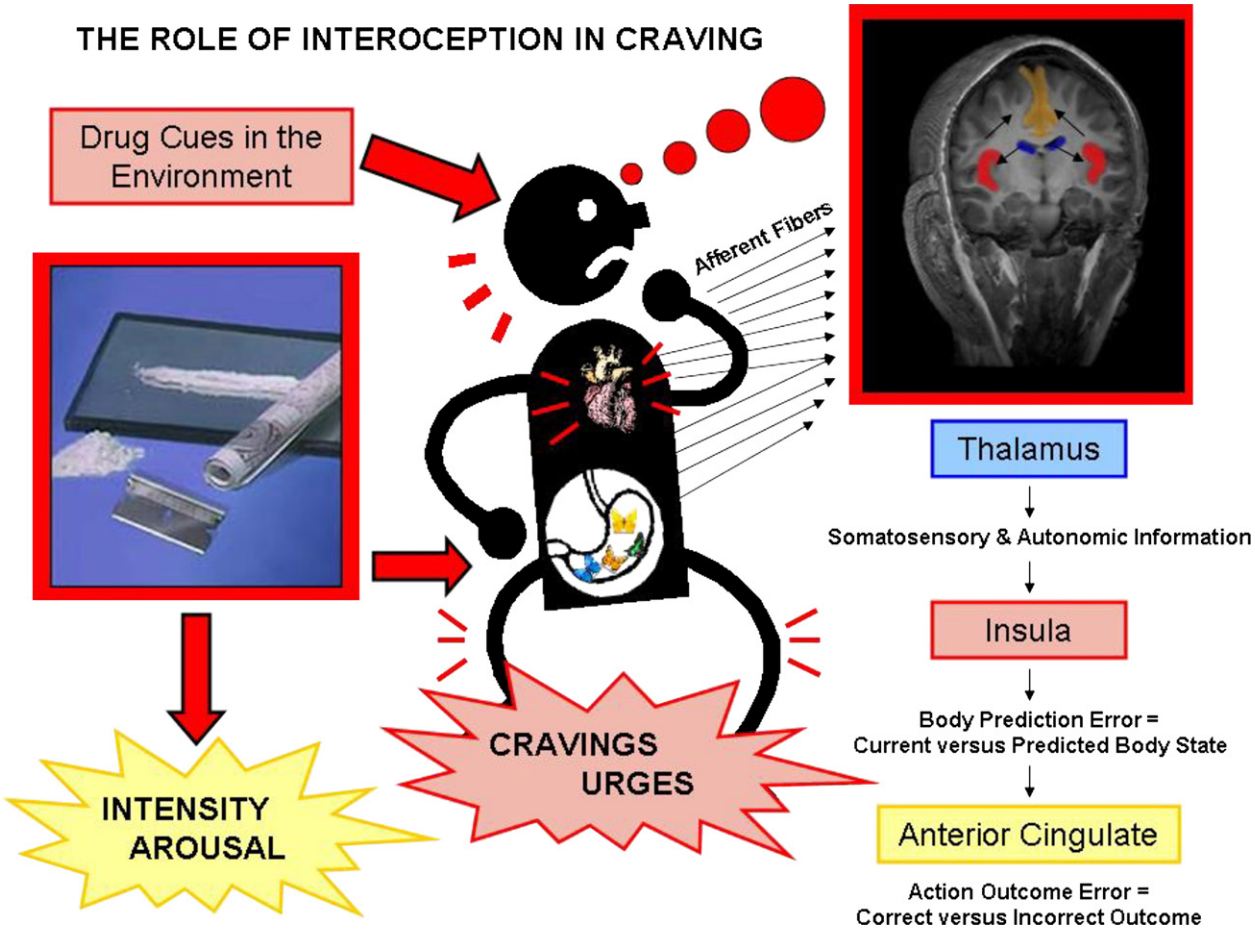


FIGURE 47.2 The role of interoception in craving: Drug cues trigger somatosensory and autonomic changes (e.g. butterflies in the stomach, racing heartbeat, sweating) indicative of heightened intensity and arousal that are relayed by afferent fibers to the thalamus (illustrated in blue) and then to the insula (shown in red), where they are synthesized and represented as self-aware feeling states associated with the craving experience. The insula generates a body prediction error that is relayed to the anterior cingulate cortex (ACC) (illustrated in yellow). The ACC then generates an error signal between the correct (or probable) versus incorrect (or improbable) potential outcome to indicate the need to regain a homeostatic balance.

paraphernalia demonstrated that left insula and the ACC correlated with cocaine craving in response to cocaine cues in cocaine abusers, the majority of whom were dependent on cocaine.

Similarly, nicotine cue-reactivity studies of craving have demonstrated results congruous with the cocaine craving literature. For instance, a PET study of nonabstinent heavy smokers and nonsmoking individuals contrasted brain activation when subjects were confronted with nicotine cues (e.g. watching a video involving a cigarette and handling a cigarette) as well as neutral stimuli (e.g. watching a nature video and handling a pencil). Changes in smokers' craving intensity from the neutral stimuli and to the nicotine cues were associated with heightened bilateral insula activation. Furthermore, an fMRI study of nicotine smokers requiring two MRI scans (one while subjects were abstinent and another when they were nicotine-satiated) showed a relationship between pre-fMRI craving for cigarettes and ACC activation during viewing of nicotine-related pictures, but only during the condition wherein smokers were required to be nicotine-abstinent for 24 h prior to their scan and were thus experiencing heightened craving during that time.

Personalized Imagery Scripts

Similarly, studies employing personalized imagery scripts of drug use also implicate insula and the ACC in craving. For example, fMRI research has shown that during guided imagery and recall of three stressful situations, left insula activation was linked to distress and cocaine craving in cocaine-dependent participants, suggesting that the insula may be linked with high negatively valenced arousal potentially exacerbated by stressful events. The interaction between craving and stress is important to examine, since stress-induced

craving states have been associated with relapse vulnerability for cocaine-dependent individuals in early recovery. Prior work in cocaine-dependent patients has demonstrated that several interoceptive sensations (racing heartbeat, bodily tension, perspiration, gastrointestinal changes, and blood flow sensations such as face flushing) were common to real-life stressful scenarios as well as situations involving explicit drug cues, and these commonalities may explain why stressful events may exacerbate the craving experience.

However, a PET study of abstinent cocaine-dependent men established that craving during cocaine scripts was correlated with left insula and the ACC compared to both neutral and anger scripts, demonstrating that activation in these regions were not due to high levels of negative affect alone. Moreover, an fMRI study examining cocaine and neutral scripts in cocaine-dependent men with and without the presence of electric shock demonstrated that higher left insula and ACC activation were present for cocaine scripts, and craving reports were the highest for the cocaine script independent of risk of electric shock, suggesting that some types of stressful situations (general threat of shock versus personalized scripts of previous stressful situations) may not activate interoceptive regions above and beyond the intensity activated by drug cues.

Drug Self-Administration Paradigms

Like cocaine cue-reactivity paradigms, a cocaine self-administration study using fMRI methodology demonstrated that ACC activation was negatively correlated with the cocaine high, but positively correlated with cocaine craving in a small sample of cocaine-dependent men. Overall, neuroimaging research of cocaine addiction provides substantial evidence that the insula (often left-lateralized) and the ACC are associated with the craving experience, presumably due to the role of the insula in predicting and monitoring an individual's current versus expected interoceptive state (e.g. detecting a homeostatic imbalance) and the role of the ACC in receiving the prediction and creating a conflict signal to initiate a change in upcoming actions and plans (e.g. obtaining and using drugs).

MODEL OF INTEROCEPTIVE DYSFUNCTION ASSOCIATED WITH CRAVING

Although the insula and ACC are crucial components of the interoceptive craving process, they function within a larger neural system to send and receive information from cortical and subcortical brain regions that serve to maintain compulsive drug use. Paulus, Tapert, and Schulteis have proposed a model of dysregulated interoception in drug addiction that explains how interoceptive processing serves to maintain drug use. When a drug-addicted individual encounters a stimulus associated with drug use, peripheral sensory receptors relay information to the thalamus regarding the location and identity of the stimulus, which is then sent to the sensory cortex and posterior insula to provide a representation of the internal body state.

In addition, centrally generated interoceptive states created from contextual memory associations are transported from the hippocampus and amygdala to the insula, which then creates an integrated current interoceptive feeling state and compares this to the predicted interoceptive state. This discrepancy is then made available to the OFC to evaluate the rewarding properties of the stimulus and the ACC to generate an error signal to initiate new motor behaviors to correct an interoceptive imbalance. Bidirectional connections between the amygdala and the striatum assist the insula in calculating this bodily prediction error and providing information on the deserved attentional salience of the stimulus. The experienced bodily state represented by the insula and the urge to act represented by the ACC are important components of the craving experience. Brain regions involved in executive control such as DLPFC then shift goal orientation in the direction of appetitive pursuit of reward propelled by craving and urges to use drugs.

SPATIOTEMPORAL DYNAMICS OF CRAVING

Comparison of Neuroimaging Techniques

Neuroimaging studies examining the spatial dynamics of craving (e.g. fMRI, PET) within the context of cocaine addiction greatly outnumber those studies using electrical measurements to determine the temporal dynamics of this phenomenon (e.g. ERPs), and it is difficult to directly compare results of hemodynamic and electrophysiological studies due to differences in spatial and temporal resolution between methods. For example, although ERP research demonstrates a link between dysfunction in lateral frontal regions and craving levels in cocaine users, due to the poor spatial resolution of EEG, it is unclear whether this dysfunction originates from the OFC, or whether it may reflect sources from the insula, the ACC, or other regions of the prefrontal cortex instead of, or in addition to the OFC unless source localization methodology is utilized in conjunction with ERP methods.

Advantages of Event-Related Potentials

ERP measures alone are helpful, however, in revealing the magnitude of attentional resources that are being recruited (reflected in ERP amplitude) and the duration that these resources are being utilized (reflected in ERP latency) to process a particular stimulus or set of stimuli. For example, two ERP studies that displayed cocaine and neutral pictures to recently abstinent cocaine-dependent individuals found that greater frontal ERP positivity was linked with cocaine craving, particularly within the context of the late positive potential (LPP), an ERP component that lasted up to 1500 ms in these studies and is thought to reflect the motivational salience of a stimulus. These findings suggest that cocaine users employ heightened attentional resources in response to cocaine stimuli for an extended duration of time, potentially due to the appetitive motivational salience that these stimuli possess, that may perpetuate the cocaine craving experience.

Although no relationship between LPP and self-report indices of craving has emerged thus far in alcohol, cannabis, or nicotine addiction research, a recent study has demonstrated this relationship within the context of opiate addiction. Within this study, abstinent heroin-dependent patients exhibited larger LPP amplitudes in central brain regions in response to heroin-related images, activation that was correlated with higher post-experiment negative reinforcement craving ratings (desired relief from negative feeling states).

Although no link between LPP and subjective craving reports has been demonstrated in alcohol or nicotine addiction, another prominent ERP component, the P300, has been related to craving in these substances. A substantial literature spanning over two decades has indicated that P300 amplitude is prominent 300–600 ms after stimulus presentation and reflects the amount of attentional resources recruited for the evaluation of stimulus significance. Thus, heightened P300 amplitude in drug users in response to drug-related stimuli may reflect heightened emotional significance and attentional bias to these stimuli. For instance, a study of abstinent current smokers, ex-smokers, and nonsmokers showed an association between heightened frontal P300 amplitude to smoking cues and ratings of the desire and intention to smoke cigarettes. Moreover, a recent study indicated that alcohol-dependent individuals exhibited higher P300 amplitude to images of alcoholic beverages than nonalcohol-dependent social drinkers, brain activation that correlated with subjective ratings of alcohol craving.

Advantages of Electroencephalography

Future research using EEG and ERP source localization methods may provide valuable information regarding the timing and sources of cortical activation associated with real-time craving processes. Source localization is a method of obtaining spatial information from EEG data wherein sources of electrical signals visible at the scalp are modeled. Although it has been assumed that the insula is activated prior to ACC during craving responses, this hypothesis has yet to be thoroughly tested, but can be evaluated using data from multielectrode EEG montages subjected to source localization. In recent years, EEG and fMRI recordings (either recorded concurrently during the same session or in separate sessions) have been used in conjunction to evaluate source and timing information, often with fMRI regions of interest utilized to determine source dipoles for initial EEG analysis.

METHODOLOGICAL ISSUES

A growing literature investigating the neural mechanisms of cocaine addiction has provided us with beneficial information regarding the brain systems involved in the craving experience. However, a few methodological issues related to self-report craving measures, sizes of samples recruited for studies, and the examination of gender differences hamper the generalizability and interpretation of these findings as a whole and can thus be improved upon.

Incorporating Self-Report Measures of Craving

It is often assumed that an experimental manipulation succeeds in inducing a craving experience without evaluating this assumption by correlating self-reported craving measures with brain activation. For example, a recent fMRI study identified tentative "craving sites" by selecting brain regions that were more active in crack cocaine users than comparison subjects during a cocaine film compared to a nature film. The researchers then examined whether any of these brain regions were more active in the users during the cocaine film compared to a sexually explicit film to rule out the possibility that the cocaine-related activations were due to general appetitive processing. Although findings demonstrated that cocaine users exhibited higher craving and greater ACC and caudate nucleus activation than during an appetitive film showing explicit sexual content, no correlations were performed between the self-reported craving index and putative "craving sites." Instead of reflecting craving processes *per se*, ACC activation in this context could instead index greater attention allocation to a conflictual, highly familiar, and emotional stimulus,

whereas dorsal striatum activation could reflect the learned experience of a cue.

Similarly, a recent fMRI study of cannabis users (most meeting criteria for cannabis dependence) asserted that marijuana cue-elicited craving was linked to greater ACC, amygdala, insula, and ventral tegmental area activation because tactile marijuana stimuli (e.g. a pipe) elicited more activation in these regions than neutral tactile stimuli (e.g. a pencil). However, no significant correlation emerged between self-reported craving ratings and these brain regions, despite the fact that a higher number of problems associated with marijuana use were linked to greater activation in brain regions implicated in reward and interoception (e.g. OFC and ACC). In studies such as this, it is crucial for researchers to link specific, reliable, and valid measurements of the craving experience (indexed by verbal reports, behavior, and/or psychophysiology) to brain activation in order to differentiate the craving experience from drug stimuli familiarity, conflict, or emotional significance that may also influence brain activation during drug-related paradigms.

Sample Size Limitations

The majority of imaging studies of cocaine craving have utilized small sample sizes (typically <20 individuals diagnosed with cocaine addiction or abuse). However, drug use disorders in general and cocaine addiction in particular is a heterogeneous disorder. Thus, different mechanisms may play a role in subgroups of cocaine-dependent individuals; for example, craving may be the result of an overwhelming ascending interoceptive afferent input, emerge from the inability to filter out drug-cue associated stimuli, or result from an inability to engage inhibitory top-down regulatory processes. Yet, small sample sizes may not allow us to detect these different mechanisms. Nevertheless, identifying the specific component processes that are dysfunctional and lead to craving could have profound clinical or treatment relevance. Small sample sizes are also common in neuroimaging studies of alcohol, cannabis, nicotine, and opiate craving, limiting generalization and power to detect significant differences as a function of the craving experience.

Examination of Gender Differences

The small sample sizes recruited for cocaine-related studies are predominately male, so it is difficult to generalize findings to women. This limitation is also true of the existing craving literatures concerning alcohol, cannabis, nicotine, and opiate addiction. The sole imaging study of gender differences in cocaine craving revealed that cocaine-dependent women exhibited greater ACC activation to cocaine scripts than neutral scripts, but cocaine-dependent men displayed greater left insula activation than women in response to cocaine scripts, suggesting that interoceptive responses to drug cues may differ in men and women, with women experiencing a more heightened error signal of homeostatic imbalance, and men experiencing more bodily intensity or arousal during interoceptive imbalance. However, in this study brain activation was not directly correlated with levels of cocaine craving, making it difficult to draw conclusions regarding gender differences in craving-related brain mechanisms.

FUTURE DIRECTIONS

Interoceptive Interventions for Drug Users

Assuming that interoceptive processing plays an integral role in the craving experience, it would be important for future investigations of drug craving to examine interventions that are targeted at processing interoceptive stimuli to alter an individual's internal state with the goal to decrease the rewarding properties of drug-related cues. For example, listening to a dripping faucet an individual experiences an increased sensitivity to exteroceptive stimuli (the auditory signal). However, one may also develop specific interventions to increase or decrease the sensitivity to interoceptive stimuli; for example, breathing, gastrointestinal, specific skin-related, or temperature stimuli, in order to decrease the urge to use related to interoceptive stimuli associated with craving.

Linking Insula Function to Positive and Negative Reinforcement

Research is also needed to examine whether intact insula and anterior cingulate function are necessary for individuals to experience conditioned positive reinforcement (approaching the rewarding aspects of the drug) or negative reinforcement (avoiding the negative consequences of drug withdrawal). Subsequent imaging studies may determine whether interoceptive craving experiences are positively or negatively valenced by requiring additional behavioral, self-report, or psychophysiological information of participants. Incorporating interoceptive processes into existing approaches to drug addiction will enhance our understanding of craving and urges within the context of drug use and relapse and point the field toward behavioral and pharmacological targets for intervention.

CONCLUSION

Neuroimaging research utilizing both hemodynamic (fMRI, PET) and electrocortical (EEG, ERP) methodologies has implicated a wide network of subcortical and cortical brain regions in the implementation of the cocaine craving experience induced by cue-reactivity, autobiographical imagery, and drug administration paradigms and measured by various self-report indices. Subcortical regions such as the amygdala, hippocampus, and striatum are implicated in conditioned reward responses that serve to perpetuate drug use, whereas frontal cortical regions that typically engage in executive control processes are instead activated by the rewarding properties of drug cues. Interoceptive sensory and autonomic information (felt as a homeostatic imbalance in response to drug cues) is relayed from the thalamus and other subcortical regions to the insula, where it is integrated into an emotional feeling state (craving) and compared to a predicted feeling state. This information is transported to the ACC, which attempts to recreate a homeostatic balance by modifying incentives to act (urges). Both the insula and ACC provide information about an individual's interoceptive state to frontal circuitry involved in executive function and subcortical circuitry involved in reward processing to maintain drug addiction. Although interoception appears to be an integral component of the craving process, additional research is warranted to determine whether interoceptive modulation in addicted individuals will result in successful new interventions.

SEE ALSO

Cocaine Addiction, Cue Reactivity, Deprivation, Craving, and Affect: Intersecting Constructs in Addiction, Craving and Expectancies, Relation of Craving and Appetitive Behavior, Defining and Assessing Drug Craving

Glossary

Amygdala a subcortical brain region located deep in the brain in close proximity to the hippocampus, which is important for processing highly salient information.

Anterior cingulate cortex (ACC) a cortical brain region that is part of the executive cognitive control system. The ACC receives interoceptive predictions from the insular cortex and generates a conflict or error signal to initiate new urges to correct an interoceptive imbalance.

Caudate nucleus a subcortical brain region, which is part of the dorsal striatum and situated near the thalamus. The caudate is important for instrumental learning and habit processing and is implicated in habitual patterns of drug urges, cravings, and use.

Cocaine a stimulant drug of abuse obtained from the leaves of the coca plant that can be orally ingested, snorted, smoked, or injected and is known to increase alertness, energy, concentration, self-esteem, euphoria, and sexuality.

Cognitive control a type of flexible information processing thought to involve the prefrontal cortex that allows behavior to change successfully as a function of current goals. Cognitive control processes include inhibition, attention allocation, and goal representation and maintenance.

Craving an emotional feeling state of strong wanting, which is accompanied by strong incentive to act, which has been implicated as an important factor in the maintenance of drug addiction and the probability of relapse.

Dextroamphetamine a stimulant drug often prescribed to treat attention deficit hyperactivity disorder that is also used recreationally as a drug of abuse. It is taken orally (in pill form) and is known for increasing and/or maintaining alertness and improving attention.

Dopamine a monoaminergic neurotransmitter associated with learning and the reward system of the brain.

Dorsal striatum a subcortical brain region comprised of the caudate nucleus and putamen.

Dorsolateral prefrontal cortex (DLPFC) a brain region situated in the prefrontal cortex thought to be involved in working memory and goal maintenance.

Drug abuse a maladaptive pattern of substance use that does not involve substance dependence but does lead to at least one symptom of clinically significant impairment or distress within a 12-month period (e.g. failure to fill major role obligations, use-related legal problems, use despite interpersonal problems, and/or use in physically hazardous situations).

Drug dependence a maladaptive pattern of substance use leading to at least three symptoms indicative of clinically significant impairment or distress within a 12-month period (e.g. tolerance, withdrawal, use in greater amounts or longer than intended, unsuccessful efforts to cut down, activities reduced due to use, significant time spent obtaining and using or recovering from use, and/or use despite exacerbation of a physical or psychological problem).

Electroencephalography (EEG) the measurement of electrical signals of the brain, usually through electrodes placed on the scalp.

Event-related potential (ERP) the average electrical signal evoked by a particular event or stimulus, such as an auditory tone or visual picture.

Exteroception the experience of intensity or arousal associated with stimuli located outside the body.

Functional magnetic resonance imaging (fMRI) a neuroimaging technique that measures the changes in blood flow and blood oxygenation associated with neural activity.

Hippocampus a subcortical brain region situated below the amygdala thought to be involved in long-term memory processing.

Homeostasis regulation of physiological processes inside the body used to maintain its internal stability and balance in response to external fluctuations.

Incentive salience the desirability of a drug that can eventually be extended to stimuli that signify or are associated with the drug (e.g. individuals associated with drug use or drug paraphernalia) that can lead to increased drug urges and cravings.

Insular cortex a cortical brain region, which is surrounded by the cortical mantle and is divided into anterior and posterior insula by the central sulcus. The insula is involved in the craving process by monitoring the current versus predicted interoceptive state of the individual and generating a signal aimed at maintaining homeostasis.

Interoception the experience of intensity or arousal associated with a homeostatic imbalance of the body, resulting in a conscious feeling of craving and the urge to use drugs.

Late positive potential (LPP) a slow-wave ERP component emerging 200 ms after presentation of a stimulus (that can last up to several seconds) thought to reflect the motivational salience of a stimulus.

Mesolimbic dopamine system a pathway carrying dopamine from the ventral tegmental area to the nucleus accumbens through the amygdala and hippocampus thought to be part of the brain's reward circuitry.

Methamphetamine a stimulant drug of abuse in the amphetamine class of drugs that can be orally ingested, snorted, smoked, or injected and is known to increase alertness, energy, concentration, self-esteem, euphoria, and sexuality.

Methylphenidate a stimulant drug primarily prescribed to treat attention deficit hyperactivity disorder that is also used recreationally as a drug of abuse. It is most commonly taken orally (in pill form) and is known for increasing and/or maintaining alertness and improving attention.

Nucleus accumbens a subcortical brain region that is part of the ventral striatum, thought to be involved in a substance rush or high as well as substance craving.

Orbitofrontal cortex (OFC) a cortical brain region located in the prefrontal cortex thought to be linked to sensitivity to reward within the context of decision making.

Parahippocampal cyrus a brain region surrounding the hippocampus thought to be involved in memory processing.

Positron emission tomography (PET) a neuroimaging technique that creates images on the basis of movement of injected radioactive material.

Prefrontal cortex the most anterior portion of the cerebral cortex that is associated with many functions including processes involved in cognitive control.

Putamen a subcortical brain region included in the dorsal striatum that is situated near the thalamus and is thought to be involved with habitual patterns of drug urges, cravings, and use.

Reticular formation a structure located in the brain stem thought to be involved in somatosensory, motor, pain, and cardiovascular modulation and control.

Source localization a computational method of determining the spatial location of electrical signals recorded at the level of the human scalp by electroencephalography.

Stimulants psychoactive drugs that result in temporary improvements of physical and/or mental function that target and stimulate the mesolimbic dopamine system. These drugs are used as prescription medications as well as drugs of abuse and dependence and include cocaine, methylphenidate, and amphetamine (and its variants such as dextroamphetamine and methamphetamine).

Thalamus a subcortical brain region located above the brain stem thought to be a gating and relay station for sensory information between subcortical and cortical areas of the brain.

Urge a strong incentive to act; for example, motivated behavior to seek, obtain, and use drugs.

Ventral striatum a subcortical brain region comprised of the nucleus accumbens.

Ventral tegmental area a subcortical brain region situated above the hippocampus that is the origin of dopaminergic neurons of the mesolimbic dopamine system.

Further Reading

Craig, A.D., 2002. How do you feel? Interoception: The sense of the physiological condition of the body. Nature Reviews Neuroscience 3, 655–666.

Goldstein, R.Z., Tomasi, D., Rajaram, S., et al., 2007. The role of the anterior cingulate and the medial orbitofrontal cortex in processing drug cues in cocaine addiction. Neuroscience 144, 1153–1159.

Gray, M.A., Critchley, H.D., 2007. Interoceptive basis to craving. Neuron 54, 183–186.

Naqvi, N.H., Bechara, A., 2009. The hidden island of addiction: The insula. Trends in Neurosciences 32, 56–67.

Paulus, M.P., Tapert, S.F., Schulteis, G., 2007. The role of interoception and alliesthesia in addiction. Pharmacology, Biochemistry and Behavior 94, 1–7.

Robinson, T.E., Berridge, K.C., 2001. Incentive-sensitization and addiction. Addiction 96, 103–114.

Verdejo-Garcia, A., Perez-Garcia, M., Bechara, A., 2006. Emotion, decision-making, and substance dependence: A somatic marker model of addiction. Current Neuropharmacology 4, 17–31.

CHAPTER

48

Interference with Concurrent Tasks

Eva Kemps and Marika Tiggemann

Flinders University, Adelaide, SA, Australia

OUTLINE

OVERVIEW

Craving can be described as a motivational state that is characterized by an intense wanting or longing to seek a particular target or activity, such as a psychoactive substance. Some of the most commonly craved substances are cigarettes, alcohol, and drugs. Although we don't normally think of food as a craved substance, we can and do in fact also crave food. When people have a craving, they often experience thoughts such as "I really need a cigarette," or "I so want a pecan sundae."

The origin of cravings has been attributed to both physiological and psychological factors. The main physiological trigger is a state of deprivation, which is accompanied by symptoms such as an increased heart rate. In the food domain, hormonal changes during the menstrual cycle and pregnancy have also been associated with cravings in women. The psychological origins of cravings include both internal and external triggers. Internal triggers can involve a range of negative emotional states, such as anxiety, depression, boredom, and stress. External triggers involve any exposure to environmental cues, such as the smell of a burning cigarette or the sight of a chocolate bar.

With the exception of food cravings, which occur among a large proportion of the general population without any problem, cravings for substances such as cigarettes, alcohol, and drugs often give rise to negative consequences. In particular, such cravings can lead to (1) development and maintenance of addictive behaviors, (2) withdrawal symptoms if consumption is resisted or prevented, and (3) relapse in people who have quit substance use. Although food cravings are not necessarily pathological, these too can be maladaptive for some individuals. They too can lead to "addictive" eating patterns and withdrawal symptoms, and certainly disrupt many people's intentions to diet. In support, research shows that food cravings have been linked to early dropout from weight loss programs. They have also been identified as a precursor to binge eating, itself a risk factor for the development of obesity and eating disorders, particularly bulimia nervosa.

While these negative consequences are what people in general might expect, there is another less obvious adverse consequence that arises from cravings. This is that cravings interfere with performance on cognitive tasks, resulting in people performing less well than they otherwise would. This chapter sets out to review the literature to date on the cognitive effects of craving. First, the chapter outlines the evidence that cravings impair cognitive task performance. Second, it considers theoretical explanations for the observed cognitive

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00048-6

interference from cravings. And finally, it discusses everyday implications of craving-induced cognitive impairments.

EVIDENCE THAT CRAVINGS INTERFERE WITH COGNITIVE PERFORMANCE

The evidence that cravings have an impact on cognition comes from two main sources: anecdotal accounts and laboratory studies.

Anecdotal evidence

Anecdotal accounts refer to what people say about themselves. These accounts consistently suggest that cravings interfere with cognitive functioning. For example, smokers have reported that their cigarette cravings interfere with their thinking. They say that when they want a cigarette badly, it takes over their thinking, and they can't concentrate. Similarly, alcoholics have complained that their cravings disrupt their daily activities by disrupting their concentration. In fact, the most consistently mentioned feature of cravings is their overwhelming nature.

Laboratory Studies

Most of the evidence that cravings impair cognition comes from laboratory studies that actually test people under controlled conditions. These studies typically induce cravings in the laboratory, and then ask people to perform standardised tests. Cravings are generally induced in research participants in one or more ways. The most common craving induction protocols involve: (1) imagery scripts, (2) substance deprivation, and (3) cue exposure.

First, imagery scripts present participants with some sort of scenario that is designed to make them crave. Participants are instructed to create an image of the described scenario. For example, cigarette smokers might be asked to imagine themselves lighting up a cigarette. A complete example of an imagery script that has been used to induce food cravings is:

> *"In a moment I am going to ask you to form an image of a scenario. It is important that you are relaxed and remain focused. Make sure you are comfortable, and clear your mind of any other thoughts. Concentrate on your breathing and try to become as relaxed as possible. There is no time limit on this exercise so take as long as you need to form a really clear and vivid image of the scenario I am about to ask you to imagine.*
>
> *Now, imagine you are eating your favorite food. Bring this experience to mind as vividly as you can, as if it were happening right now. Provide as much detail from your imagination as you possibly can to make the image as real as if you are actually eating the food. When you feel you have formed the clearest and best image of this scenario that you can, let me know."*

Second, substance deprivation involves instructing participants to abstain from the craved substance for a set period of time before the laboratory session. Usually participants are deprived for 6, 12, or 24 h. So, cigarette smokers might be asked to not smoke for 12 h. Or, people who crave chocolate might be asked not to eat it for a day.

Third, cue exposure involves exposing participants to the actual substance in the laboratory (but without allowing them to consume it). For example, in studies of alcohol craving, to induce a craving, participants might be presented with a glass containing an alcoholic beverage. Similarly, in the food domain, chocolate cravings might be induced by showing participants actual chocolate.

Following one or more of these induction procedures, participants are generally asked to report their level of craving. This is generally done by completing a rating scale that ranges from "no urge at all" to "extremely strong urge to smoke, drink, eat chocolate (as appropriate)." Each of the three urge induction protocols (imagery scripts, substance deprivation, cue exposure) has been shown to individually successfully elicit cravings in the laboratory. But the most effective methodology for eliciting cravings seems to be substance deprivation in conjunction with cue exposure. For example, smokers who abstain from smoking for 12 h and are subsequently exposed to a lit cigarette in the laboratory report stronger cigarette cravings than those who are only nicotine deprived or are only exposed to cigarettes. Additionally, cue exposure is maximally effective if the participant interacts with the craving induction stimulus in some way. For example, asking a smoker to pretend that they are smoking a lit cigarette that they are holding produces stronger cravings than asking them to hold the lit cigarette, which produces stronger cravings than asking them simply to view a lit cigarette.

Next, participants are asked to perform a cognitive task. The earliest laboratory studies have used a simple reaction time task to examine the cognitive effects of craving. In this task, participants are presented with a series of stimuli. These are presented either visually (e.g. a series of asterisks that appear on a computer screen) or orally (e.g. a series of tones). The stimuli are presented at varying and random intervals. Participants have to respond (by pressing a button) as quickly as possible whenever they see or hear the stimulus.

Using this task, a number of studies have shown that experimentally induced cigarette cravings increase reaction times in smokers. For example, in one study regular smokers created images of scenarios. Some were about smoking (e.g. smoking at a party with friends) and others were not about smoking (e.g. watching and admiring a fireworks display). The participants were asked to imagine themselves participating in these scenarios and subsequently rated their urge to smoke. They then performed a reaction time task in which they were instructed to press a button as quickly as possible whenever they heard a tone. Smokers who imagined smoking scenarios reported stronger cigarette cravings and took longer to respond to the series of tones than those who imagined nonsmoking scenarios. Similar increases in response times have also been shown in alcoholics, opiate users, and weight loss dieters following experimental induction of alcohol, drug, and food cravings, respectively.

Cognitive decrements during craving episodes are not limited to slower response times to stimuli on a simple reaction time task, but also occur on other, more complex tasks. In particular, several studies have shown that experimentally induced cigarette cravings impair language comprehension and mental arithmetic in smokers. In each study, cigarette cravings were induced via an imagery script. Half the participants were asked to form an image of a scenario about smoking (i.e. smoking after a satisfying meal in a restaurant). The other half imagined a scenario that was in no way related to smoking (i.e. relaxing and enjoying the view from a window). For the language comprehension task, participants read sentences and answered true–false questions about them. The mental arithmetic task involved participants verifying complex addition problems. Participants who imagined the smoking scenario had stronger cigarette cravings and made more errors on both the sentence comprehension and mental addition tasks than those who imagined the nonsmoking scenario.

In another study, habitual chocolate cravers showed performance decrements on a working memory capacity task following a chocolate urge induction. Working memory capacity refers to the ability to actively store information while simultaneously processing other information, and was assessed by the operation span task. In this task, participants were required to remember a series of words (storage) while also verifying mathematical equations (processing). Chocolate cravings were elicited through a combination of deprivation and exposure to actual chocolate. Specifically, a random half of participants (craving group) were instructed not to eat chocolate for 24 h, and upon arriving at the laboratory, were presented with a selection of wrapped chocolate bars attractively presented in a basket. Participants were asked to choose their favorite chocolate bar, to unwrap it and place it and the wrapper on a small tray positioned in view below the computer monitor, and to indicate how much they liked the chocolate. Participants then performed the operation span task in the presence of the chocolate bars. The other half of the participants (control group) did not abstain from eating chocolate and were instead presented with a basket of similar sized colored wooden blocks and analogously asked to choose their favorite colored block and place it in view below the computer monitor. The results showed that participants in the craving group reported stronger chocolate cravings than those in the control group. More importantly, as predicted, they also both recalled fewer words, and showed slower mathematical equation verification times in the operation span task.

Why Do Cravings Disrupt Cognitive Performance?

Taken together, the findings from laboratory studies clearly show that cravings impair performance on a range of cognitive tasks, from simple (like reaction time) to complex tasks (like language comprehension). There are a number of different theoretical explanations for these adverse cognitive effects of craving. The most supported is that dealing with cravings requires mental effort. In other words, cravings take up cognitive capacity. Because our cognitive capacity is limited, there is then less available for competing cognitive demands, such as a reaction time task. Consequently, task performance suffers. In this way, the interference of cravings with concurrent cognitive tasks can be likened to a dual-task procedure. In a dual-task procedure, participants are asked to perform two tasks simultaneously. The aim is to determine the amount of cognitive capacity that is used by the secondary task (e.g. counting numbers) by measuring how much it disrupts performance on the primary task (e.g. proof reading). Thus, the greater the performance decrement on the primary task (proof reading), the more capacity is being consumed by the secondary task (counting numbers). As such, cravings behave like a secondary task, taking up limited capacity that is required for the performance of a concurrent cognitive task.

Theoretical accounts of craving suggest two possible general mechanisms for why cravings take up cognitive capacity. First, during craving episodes, individuals may direct cognitive capacity to processing craving-related cues. In other words, their attention is automatically drawn to these cues, and they are distracted by them. For example, they notice the cigarette advertisement on television, or someone smoking, or the smell of coffee. Or, alternatively, individuals may allocate

cognitive capacity toward inhibiting consumption behavior. Consumption behavior, such as smoking and eating, is itself automatic in the sense that it does not take up cognitive capacity. However, once such behavior is triggered, for instance by craving, cognitive capacity is required to inhibit it. Thus, individuals need to consciously try to stop themselves from smoking, or drinking, or eating, and this takes up cognitive capacity.

Cravings Interfere Specifically with Visual Tasks

As mentioned previously, instructing participants to create an image of a craved substance is an effective means of eliciting cravings in the laboratory. Converging evidence from a number of recent studies in fact points to mental imagery, particularly visual mental imagery, as playing a key role in cravings. First, anecdotal reports of naturally occurring craving experiences show that when people crave, they have vivid visual images of the craved substance. For example, smokers commonly experience images of "relaxing with a cigarette" or "having a cigarette to help cope with stress." Second, several studies have reported a positive correlation between the vividness of people's images and the intensity of their craving, such that stronger cravings are associated with more vivid images. Third, surveys of everyday cravings for various substances including alcohol, tobacco, and food show that respondents readily use imagery terms to describe their cravings. For example, in one study, respondents used phrases such as "I could picture the pizza in my mind, picture eating it" when asked to write a short paragraph describing a previous food craving episode. In addition, when presented with a list of descriptive statements, respondents strongly endorsed imagery-based descriptors as characteristic of their cravings. Imagery descriptors in the visual modality (e.g. "I am visualizing the food") in particular were rated highly. In contrast, auditory descriptors (e.g. "I can hear myself having the food") were not highly rated. Together, these lines of evidence led to the conclusion that mental imagery is a key element of cravings. They further indicate that the imagery component of cravings is predominantly visual rather than auditory or verbal in nature.

We know from more general cognitive psychological research that visual imagery interferes selectively with visual tasks. For example, holding a visual image in mind interferes specifically with the detection of visual signals, but does not interfere so much with an auditory task, such as detecting tones. Conversely, an auditory image will interfere more with an auditory task than a visual task. This modality-specific interference occurs because the visual (and auditory) system has limited capacity. Therefore, a visual secondary task interferes with a visual primary task because they both compete for the same limited capacity visual system.

In view of the visual imagery basis of cravings, it logically follows that cravings will interfere most with performance on a visual task. This was demonstrated in a study that specifically examined the impact of chocolate cravings on both a visual memory task and a comparison verbal memory task. Half the participants (craving group) underwent a chocolate craving induction protocol that involved a combination of deprivation and cue exposure. Specifically, participants abstained from eating chocolate for 24 h before the testing session. They then chose their favorite from a selection of fun-size wrapped chocolate bars, unwrapped it, and performed the memory tasks in the presence of chocolate. The other half (control group) were not deprived of, nor exposed to, chocolate. The verbal memory task consisted of the digit span task. In this task, participants are presented with random series of digits and asked to recall these in correct serial order. The visual memory task consisted of the Corsi blocks task. In this task, the experimenter points to a random selection of nine blocks irregularly positioned on a board; the participant's task is to repeat the presented sequence. The Corsi blocks task has been designed to be analogous to the digit span task, but in the visual domain. As predicted on the basis of the visual imagery nature of cravings, participants in the craving condition performed more poorly on the Corsi blocks task than control participants, but the groups did not differ on the digit span task. These results indicate that cravings selectively disrupt performance on visual tasks.

Real-World Implications

It needs to be noted that the cognitive effects of individual craving episodes are likely to be small. For example, across craving studies for different substances, responses on a simple reaction time task are on average 50–100 ms slower when people are craving than when they are not. In addition, accuracy in reading comprehension and mental arithmetic in smokers drops by about 10% when they crave a cigarette. Likewise, memory performance drops by about 5–10% when people crave chocolate.

Nevertheless, even small cognitive decrements have the potential to compromise the ability of people to perform their jobs and everyday cognitive tasks to an optimal level, particularly as such decrements are likely to be cumulative over time. For example, most work places are now smoke-free, so smokers cannot respond to their smoking urges while at work. But they do often have to process complex information as part of their work. Not being able to respond to cigarette cravings in this situation may reduce smokers' work efficiency.

Similarly, there is an abundance of food and eating cues (e.g. fast food outlets, television advertisements) in our contemporary environment that are likely to induce food cravings in vulnerable individuals. Such individuals will often be exposed to craving cues in situations in which they have to perform a cognitive task but cannot respond to their food craving, for example, a driver who sees an advertisement for a chocolate bar on a billboard.

More generally, there are situations where split-second responses are crucial, for example, when maneuvering through dense traffic or responding to a visual stimulus in real-world vigilance tasks, such as monitoring a radar screen or inspecting items on an industrial production line. Unsatisfied cravings in any of these situations have the potential to increase accidents.

CONCLUSION

Cravings can give rise to a range of negative consequences, including cognitive impairment. Research to date shows that cravings for cigarettes, alcohol, drugs, as well as food disrupt performance on a range of cognitive tasks, including reaction time, memory, language processing and mental arithmetic; thus people do not perform to their maximum. These adverse cognitive effects further highlight the need for interventions designed to reduce the occurrence of cravings.

SEE ALSO

Cue Reactivity, Implicit and Associative Processes in Addiction, Sensory Imagery in Craving, Attentional Biases in Craving

Glossary

Binge eating the consumption of an excessively large amount of food in a single sitting, accompanied by feelings of lacking control over one's eating

Bulimia nervosa an eating disorder characterized by recurrent binge eating, followed by compensatory behaviors, such as fasting, excessive exercising, and vomiting

Cognitive capacity the amount of mental resources available for processing information

Cognitive impairment substandard performance on mental ability tasks

Craving a strong urge or desire to consume a particular substance

Mental imagery the mental representation of stimuli that are not physically present

Obesity a medical condition characterized by excessive body fat, which can lead to health risks, such as heart disease, high blood pressure, and diabetes

Working memory capacity the ability to temporarily store information in the face of ongoing processing

Further Reading

Baxter, B.W., Hinson, R.E., 2001. Is smoking automatic? Demands of smoking behaviour on attentional resources. Journal of Abnormal Psychology 110, 59–66.

Cepeda-Benito, A., Tiffany, S.T., 1996. The use of a dual-task procedure for the assessment of cognitive effort associated with cigarette craving. Psychopharmacology 127, 155–163.

Hillebrand, J., 2000. New perspectives on the manipulation of opiate urges and the assessment of cognitive effort associate with opiate urges. Addictive Behaviors 25, 139–143.

Kemps, E., Tiggemann, M., Grigg, M., 2008. Food cravings consume limited cognitive resources. Journal of Experimental Psychology: Applied 14, 247–254.

Madden, C.J., Zwaan, R.A., 2001. The impact of smoking urges on working memory performance. Experimental and Clinical Psychopharmacology 9, 418–424.

Sayette, M.A., Hufford, M.R., 1994. Effects of cue exposure and deprivation on cognitive resources in smokers. Journal of Abnormal Psychology 103, 812–818.

Sayette, M.A., Monti, P.M., Rohsenow, D.J., et al., 1994. The effects of cue exposure on attention in male alcoholics. Journal of Studies on Alcohol 55, 629–634.

Tiggemann, M., Kemps, E., Parnell, J., 2010. The selective impact of chocolate craving on visuospatial working memory. Appetite 55, 44–48.

Zwaan, R.A., Stanfield, R.A., Madden, C.J., 2000. How persistent is the effect of smoking urges on cognitive performance? Experimental and Clinical Psychopharmacology 8, 518–523.

Zwaan, R.A., Truitt, T.P., 1998. Smoking urges affect language processing. Experimental and Clinical Psychopharmacology 6, 325–330.

Relevant Websites

http://www.psychologicalscience.org/index.php/news/releases/the-psychology-of-food-cravings.html – Association for Psychological Science press release

http://www.dailymail.co.uk/sciencetech/article-1279425/Food-cravings-bad-brain-waistline.html – Daily Mail, Science and Technology article

CHAPTER

49

Relation of Craving and Appetitive Behavior

Stuart G. Ferguson[*,$], *Saul Shiffman*[**]

[*]School of Pharmacy, University of Tasmania, Hobart, TAS, Australia, [$]Menzies Research Institute Tasmania, University of Tasmania, Hobart, TAS, Australia, [**]University of Pittsburgh, Pittsburgh, PA, USA

OUTLINE

The phenomenon of craving has been a core focus of dependence research for many decades. The push to understand and treat craving is born out of the prominence of craving in addicts' experience of their addiction, and the wealth of evidence that suggests that the experience of craving can motivate behaviors such as smoking, gambling, or eating, even in the presence of a conscious desire to avoid the behavior in question. The ability of cravings to induce or promote behaviors, seemingly against the initial intention of those who experience them, is the core characteristic of craving and the one that most attracts the attention of researchers and theoreticians. We also discuss some of the reasons why some studies may have failed to find a relationship between craving and behavior.

Arguably the most intensive investigations on the nature and function of craving have taken place in the context of drug dependence, the findings from which are the focus of this review. As we discuss in detail below, across substances of dependence, drug craving has consistently been found to motivate drug-seeking behaviors and, hence, to disrupt efforts by users to achieve and maintain abstinence. Below we explore the nature and experience of craving, starting with the definition and measurement of craving. The evidence relating to the role of craving in maintaining appetitive behaviors is then examined. Although we have chosen to focus on findings from the drug dependency literature, craving has also been documented in relation to appetitive behaviors that are not substance related (e.g. gambling). The principles and knowledge gleaned from studies of substance abuse are also likely to be applicable to the experience of craving in these fields.

CRAVING: DEFINITION AND CONCEPTUALIZATION

For a topic that has been the focus of such a wealth of scientific research over the last century across a broad range of behaviors – most notably in drug dependence – it is perhaps surprising to learn that scientists are yet to categorically define what is meant by the term craving. Across, and sometimes even within disciplines, the term craving has been defined and operationalized in many ways. Although not identical, the many available definitions of craving have core similarities; even though the words may be different, the individual definitions contain similar themes. Thus, rather than attempting to define the concept of craving here, we instead focus on describing these core overarching

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00049-8

concepts that appear to span the individual definitions that other researchers have utilized.

It is generally agreed that a craving is primarily a relatively intense need, urge, or strong desire to undertake a target behavior. Cravings vary in intensity, but at their most intense, are often described by those who experience them as being overwhelming, provoking to some degree a loss of control over one's reactions. Many definitions of craving often incorporate some component of perceived loss of control over one's behavior. One may desire an ice cream, for example, but unlike a craving, such desires do not generally rise to the level where they are disruptive to other facets of life.

Conceptually, craving is often thought of as a withdrawal effect; that is, simply a product of abstinence from a target behavior or substance. For example, a typical physiological model of drug craving would posit that dependent individuals have become tolerant to the presence of a drug, resulting in changes in the very physiology of the brain (e.g. through changes in the number and/or sensitivity of receptors) and, as a result of these changes, they experience craving when the drug is not available in sufficient quantities. Within such a framework, craving is thought to follow a similar time course to withdrawal symptoms during periods of abstinence. Typically, drug withdrawal symptoms increase with the duration of abstinence before peaking and then gradually returning to baseline levels (although researchers have identified a number of unique time profiles of withdrawal symptoms that do not follow this trajectory). Supporting such a conceptualization, there is strong evidence from numerous studies across many drugs of dependence that the experience of craving does vary predictably with abstinence (typically increasing within the hours following cessation of drug use before tapering back to baseline levels gradually over subsequent days and weeks).

However, unlike true withdrawal effects, researchers have also noticed that craving can be induced without manipulating abstinence; researchers have demonstrated that craving can be reliably invoked even during periods of drug satiation or after prolonged abstinence (when withdrawal effects should have disappeared). Using laboratory-based cue reactivity paradigms (discussed briefly below and in detail in other chapters), for example, researchers have demonstrated that intense cravings can be induced simply by exposing drug users to cues (or stimuli) that are associated with drug administration (e.g. a syringe or drug paraphernalia). Certain emotional states, such as emotional distress and negative affect, have also been shown to reliably provoke craving. Such moment-to-moment volatility in craving is difficult to adequately explain using a purely physiological/withdrawal-based model of craving: if abstinence were the only factor driving craving, then the experience of craving should follow a reliable, relatively consistent, and slow time course (as circulating levels of the drugs are cleared from a user's system). A further challenge to traditional withdrawal-based explanations of craving is the observation that craving continues to fluctuate even after circulating levels of the drug are completely exhausted. Withdrawal-based explanations of craving cannot explain why individuals can continue to experience fluctuations in craving months or even years after complete cessation was achieved.

Such findings beg the question: How can an intense craving for a drug be induced simply via exposure to certain stimuli or cue? The answer, most theorists argue, lies in the psychological concept of conditioned associations. Within this conceptual framework, the link between situational stimuli and craving is thought to be mediated by conditioned, or learned, association; that is, these cues have come to elicit craving through a history of repeated pairing with smoking. There are various accounts for how such relationships are established, but they all stress the importance of repeated pairing of stimuli with drug administration.

Thus, unlike traditional withdrawal symptoms, craving appears to be driven not only by abstinence, but also by exposure to drug-related stimuli. For this reason, researchers have found it useful to distinguish between two types of craving depending on how it is induced, namely abstinence-induced craving and cue-induced cravings. (Cues in this context may be external (sensory stimuli such as visual or olfactory reminders of a drug) or internal (strong emotions)). Abstinence-induced craving, also known as background craving, is akin to withdrawal symptoms and is characterized by its relationship to drug levels. Such craving is experienced as relatively steady, tonic states over the course of days or at least hours, fluctuating relatively slowly. Cue-induced cravings, on the other hand, are more volatile and subject to moment-to-moment variations in intensity. The example of hunger can be used to illustrate the distinction between these two types or conceptualizations of craving. If one has not eaten, one is likely to feel constantly hungry. Such hunger is akin to background craving (or withdrawal symptoms); craving that changes slowly over the course of days and or weeks depending on an individual's degree of abstinence. In contrast to the relatively steady experience of background craving, however, as discussed above, drug users also experience bouts of intense craving; episodic spikes or surges in craving that are seemingly independent of abstinence. Such cravings have been referred to as cue-induced cravings. Using the hunger analogy, consider the desire for food one experiences when seeing or smelling appetizing food. The mere presentation of food can result in intense, acute, hunger even when one is sated. In such circumstances, the

person is not hungry because they require sustenance in the traditional sense; rather they experience the sensation of hunger because they have seen that food is available.

Cue-induced cravings are experienced as acute episodes of intense craving overlaid on slowly varying levels of background craving. As discussed below, however, the measurement tools used to assess craving do not differ according to the type of craving being measured.

ASSESSING CRAVING

Unlike the measurement of physical properties (e.g. weight, blood pressure, heart rate, etc.) psychological concepts such as craving cannot be measured directly. Researchers have to either ask patients to self-report their levels of craving or measure secondary physiological properties that are believed to vary systematically with a patient's level of craving intensity (proxy measures of craving). As discussed briefly (and in more detail in other chapters), there are advantages and disadvantages to each of these approaches.

Craving is most commonly assessed through individual self-report measures. Patients are explicitly asked to report on their subjective level or intensity of craving. Such questioning is typically conducted using questionnaires, but it is also possible to have patients self-report their craving verbally or by turning a dial. Several forms of self-report assessments have been used in craving research. These include verbal reports, visual analog scales (individuals are asked to indicate their degree, or amount, of agreement with a statement by marking a position along a line between two anchors or end points), Likert-type scales (rating scales), and magnitude estimation scales, among others. Such measures range from single-item reports (e.g. "How much are you craving right now?") to multi-item questionnaires that use a variety of items to capture different aspects of craving, and either average across these items or produce a multifactorial measure of craving. As indicated by the range of different approaches, there is no universally accepted measure of craving, and the type of assessment used may be contingent on various factors, including study design, outcome of interest, and a number of other factors. Although relatively simple to use, self-report measures of craving are not without their problems, the most serious of which is that self-report measures are inherently subjective. It is impossible to know whether one person's experience of craving is the same as the next. For this reason, researchers have explored alternate ways to assess the experience of craving in the hope of developing more objective measures.

In an attempt to overcome the issues involved with self-reporting, researchers have developed many ostensibly more objective assessments of craving. These include psychophysiological measures, cognitive processing and performance (where craving is implied by decrements in cognitive speed and/or accuracy), behavioral indices, measures of facial expressions, startle response, and reinforcement proxies (e.g. delay discounting).

More recently, researchers have started to investigate whether craving can be reliably indexed using neuroimaging techniques such as functional magnetic resonance imaging. In such studies, researchers attempt to assess a patient's intensity of craving using patterns of activity in various regions of the brain. The brain areas most commonly activated across studies of cue-provoked craving include those associated with executive control, and emotional and cognitive learning and processing. Identification of these areas is important not only to understand the neural underpinnings of craving but also to begin to predict, understand, and perhaps eventually manipulate the neural underpinnings of successful regulation of craving. In the case of appetitive behaviors, the ability to regulate craving is likely vital to forestall relapse, lending importance to the investigation of such regulatory processes. Such a neurocognitive model of self-regulation seeks to understand the interactions between brain regions that have been identified as being important to the subjective experience of craving and can be applied to the study of either withdrawal or cue-induced craving. Of particular interest, findings from brain imagining studies across a range of dependence-producing substances suggest that cravings may be the same (or at least similar) regardless of the stimulus that provoked them. Across substances, similar patterns of brain activity are observed under conditions that generate cue-induced cravings, suggesting a shared circuitry for craving across substances. If this is indeed the case, it suggests the tantalizing possibility of the development of generic (i.e. behavior nonspecific) treatments for the successful regulation of craving in the future.

Such non–self-report measures of craving are typically viewed as correlates of craving or peripheral indicators of an individual's subjective experience of craving, meaning that although they may appear to be more objective, they are actually more distally related to the experience of craving and hence potentially poorer measures of the actual craving experience. For example, it may be difficult to know if brain activation in response to a stimulus in the laboratory is reflective of craving, the process of regulation of that craving, or a combination of the two. These methodological issues may become more pronounced during certain conditions (e.g. if the substance in question is readily available, thus tending to engender greater experience of craving) and in certain

subject populations (e.g. in abstinence-seeking substance users, who may tend to resist their craving). A further issue with using objective proxy measures of craving is that these measures are not uniquely related to craving. The objective measures noted above are all subject to variation that is unrelated to an individual's experience of craving; thus, it is difficult to tease apart whether observed fluctuations in these measures are true indications of changes in craving, or whether they are simply indexing variation in unrelated phenomena. Nevertheless, despite these limitations, given that self-report measures of craving are by no means a perfect index of the craving construct, such additional indirect measures may be viewed as interrelated components that comprise the phenomenon of craving.

The type of measure used (self-report or physiological) traditionally does not differ depending on the way that the experience of craving has been induced. Such an approach is born out of the belief (explicit or otherwise) that how craving is induced, be it via abstinence or via cue exposure, does not fundamentally change the nature of the craving itself. To put it more simply: all cravings are fundamentally the same regardless of how they are generated. However, the assumption that abstinence-induced and cue-induced cravings are simply just the same phenomenon induced through different mechanisms has not been extensively tested, although as noted earlier, recent brain research does suggest similarities to craving across substances of dependence, which is suggestive of a common experience of craving.

THE STUDY OF CRAVING: LABORATORY AND FIELD-BASED RESEARCH METHODOLOGIES

To truly understand craving, researchers need to develop methods to apply those tools in real-life situations where individuals experience craving in real time. While this can be, and indeed has been, conducted in real-world situations (i.e. in observational field studies), most of our understanding of the intensity, duration, and motivational significance of craving (i.e. its relationship to behavioral change) comes from detailed observations taken in laboratory settings.

Craving has been studied extensively in laboratory settings for decades. Such studies are generally referred to as cue reactivity studies because they traditionally utilize situational cues, or stimuli, to artificially invoke intense cravings under controlled conditions. The standard cue reactivity research paradigm is as follows: a participant undergoes a baseline assessment, which may include both psychological (e.g. cigarette craving before cue exposure, mood rating) and/or objective physiological (e.g. heart rate) measures. Next, the participant is exposed to a cue or stimulus that has been designed to induce craving. During and/or following this cue exposure period, the participant then again completes a battery of assessments similar to those administered at baseline. Differences in measures before and after exposure to the target cues are then used to calculate the magnitude of the craving induced, based on the assumption that exposure to the cues will influence the individual's level of craving. Exposure to a neutral cue may also be used in a separate session as a control.

The actual cue or stimulus used to induce craving obviously varies depending on the nature of the craving that one is trying to induce. If the participants are nicotine dependent, for example, the cue may be a lit cigarette. However, if the participants suffer from a gambling disorder, then researchers may choose to use a gambling-specific cue, such as poker chips or a pack of cards. Furthermore, researchers can choose whether to use physical or nonphysical stimuli. Examples of nonphysical exposure to stimuli that have been used include visual cues (e.g. showing participants pictures of relevant stimuli) or imagery (having participants imagine that they are in the presence of relevant stimuli). Some more recent laboratory studies have used virtual reality in an attempt to recreate the more complex environmental stimuli that individuals encounter in the real world (i.e. outside the laboratory setting). In general, nonphysical stimuli have been found to yield less robust increases in craving than physical ones, perhaps because physical stimuli are richer more detailed stimuli, but in general they are easier to use and make it possible to expose participants to a broader range of stimuli. In an attempt to make them more salient and hence more potent, some studies that utilize nonphysical cues have attempted to personalize these cues such that places or things that individuals identify as cues in their everyday lives can be incorporated into cue presentations within the laboratory. As discussed earlier, researchers have used two primary mechanisms for inducing craving in individuals: abstinence and cue exposure. Although cue exposure is the most commonly used laboratory cue induction strategy implemented, abstinence has also been used in some studies of craving using similar designs. The study of abstinence-induced cravings in laboratory settings, however, is less common due, in part, to the practical complications of conducting such studies. In the case of drug studies, for example, an abstinence induction procedure requires recruiting drug dependent individuals and then asking them to abstain from drug administration for a set duration of time; a procedure that is both uncomfortable for the participant and potentially time-consuming for the experimenter.

Cue reactivity studies have allowed researchers to closely explore topics such as the role of cues in

provoking craving, but their artificial nature means that it is difficult to know how well these findings generalize to real-world experiences. Furthermore, cue reactivity studies are limited practically in terms of the duration that craving can be monitored, meaning that they are generally unsuitable for the study of the extended time course of craving, or the effect of these patterns on behavior. For these reasons, researchers have also attempted to study such craving outside the laboratory, in natural real-world settings. Such longitudinal observation studies have the advantage of assessing the relationship between craving and target behaviors within the context in which they naturally occur; that is, within an individual's general day-to-day life.

One relatively recent development in the field-based study of craving has been the use of real-time measurement (e.g. ecological momentary assessment (EMA)). This approach avoids retrospective recall by having participants report their experience of craving as it is happening. Participants may be instructed to report their craving at particular milestones or events (e.g. when they are about to use a drug or about to relapse after changing their behavior), or their craving may be monitored by randomly sampling their craving state over the course of the waking day (e.g. by beeping them at random times to ask about their craving). This type of field-based real-time data allows for collection of rich data over time, in individuals' natural environments.

As discussed earlier, some cravings are believed to be induced through exposure to situational stimuli. Although this can and has been studied within a laboratory setting, in order to truly understand the behavioral significance of such craving, we need to observe them, and their consequences, under real-world conditions. However, accurately monitoring craving within a real-world setting is more complicated than doing so in a laboratory. Perhaps the most common method of assessing craving within a participant's natural environment is to use retrospective recall: that is, participants report how much craving they experienced, and/or how frequently they experienced craving, over a given period of time (e.g. the previous week). Asking a participant to report how they felt at some previous point in time, however, involves the process of recall: participants need to think back over the defined period to remember how much craving they experienced. Such an approach is problematic because recall has been found to be both fallible (i.e. people forget) and potentially biased (i.e. systematically incorrect).

In an attempt to overcome some of these issues, researchers have utilized diary methods, where participants complete daily (or more frequent) assessments of their craving. More sophisticated methodologies utilize handheld computers to administer surveys of craving in near real-time as described above, further removing the opportunity for recall bias to influence the study results.

Regardless of the data collection strategy used, field-based monitoring of craving allows for the measurement of the natural history of craving and the study of temporal patterns. Together with results gleaned from cue reactivity studies, these studies have helped researchers to better understand the relationship between craving and appetitive behaviors. An overview of the results from these studies is the focus of the next section.

RELATIONSHIP BETWEEN CRAVING AND BEHAVIOR

As discussed earlier, although the proposed mechanism of action differs, all theories of craving to some extent acknowledge their role in motivating, or driving, appetitive behaviors. Clinically, across drugs of dependence and appetitive behaviors in general, craving is one of the symptoms that most concerns dependent individuals who are interested in quitting and, understandably, is one of the symptoms that they commonly seek to ameliorate through treatment. A cornerstone of treatment for appetitive behavior is aiding individuals to strengthen their own ability to regulate craving during high-risk situations, such as stressful situations or situations involving relevant cue exposure. Typically, reaching this overarching treatment goal may be achieved through a combination of pharmacologic and behavioral therapy. From a research perspective, a better understanding of the process of craving – and self-regulation in the face of craving – is central to improving health across disease states. Thus, craving is an exemplary condition for study, because unchecked, it often leads to behavior and outcomes that are against a person's interest and often contrary to their plans (e.g. cigarette craving during a quit attempt may lead to relapse and the universally undesirable high risk of multiple diseases and or markedly premature death).

Findings from both observation field studies and laboratory-based studies suggest that a drug-dependent individual's concern regarding the experience of craving is well founded. Extant research suggests that increases in craving are associated with increased likelihood of engaging in a given behavior in real time. For example, real-time data collected via handheld computers have shown that day-to-day variation in craving during a smoking cessation attempt accurately predicts the likelihood of lapsing. There is some evidence that craving predicts cessation success even when the severity of withdrawal symptoms (such as negative affect) is accounted for; suggesting that craving is more than just simple drug withdrawal. In addition, during conscious

efforts to abstain, individuals often cite increased craving as a significant contributor to resuming use of the drug.

The relationship between craving and drug use is seen not only in observation studies but also under controlled laboratory conditions. For example, cue reactivity studies have shown that artificially induced increases in craving are associated with subsequent drug use in individuals not currently attempting to abstain.

Another source of evidence supporting the relationship between craving and behavior can be found in the success of pharmacotherapies that target, or purport to target, craving. A prime example of this is nicotine replacement therapy (NRT) for the treatment of nicotine dependence. NRT was developed to mitigate the loss of nicotine that occurs during cessation attempts, thereby reducing withdrawal symptom severity and craving and making abstinence more sustainable. Since its development, numerous studies have shown that NRT promotes abstinence and it reduces craving severity during a cessation attempt. The one study that directly tested whether the reduction in craving and withdrawal symptom severity did indeed account for the effect of NRT on cessation found some support for the proposed mechanism of action: although incomplete, NRT's ability to promote cessation was found to be partially mediated via its effects on craving. The fact that medications that aim to reduce craving have been found to promote cessation is further evidence of the relationship between craving and appetitive behavior.

This is not to say that the relationship between craving and behavior has always been clearly supported. Some field-based and laboratory studies have failed to show the hypothesized relationship. However, a number of methodological issues have been proposed that may explain some of these inconsistent findings. Being an abstract concept, craving is difficult to accurately measure, particularly in real-world situations. A growing number of studies have utilized EMA real-time data collection methods described above, but the majority of field-based evidence has relied on retrospective recall and hence is susceptible to recall bias and/or error, perhaps clouding a legitimate relationship between craving and appetitive behaviors.

In addition to the difficulties involved in the assessment of craving itself, a second methodological issue that faces researchers interested in exploring the role of craving in appetitive behaviors has been first identifying, and then controlling for, factors that may moderate the effect of craving. A good example of this is the effective use of coping strategies. If unmeasured, the successful use of strategies to resist craving can obscure the relationship, because they potentially allow high levels of craving to occur without increasing the probability of the target behavior occurring; the effect of craving on behavior is moderated by a third variable, in this case the use of coping mechanisms. Failure to first assess, and then account for, such moderating variables can account for some of the negative findings in this area.

A final methodological challenge for researchers is that there is some evidence that the relationship between craving intensity and behavior may not be completely linear in nature. There is evidence of a positive linear relationship between craving and the probability of drug use at lower levels of craving intensity but, once craving reaches a certain intensity or threshold, the relationship asymptotes, with further increases in craving not increasing the probability of drug use. The exact intensity of this threshold likely differs from person to person, and indeed likely even from situation to situation within the same individual. One important implication of this nonlinear response curve is that a relationship between craving and use will only be seen at low to moderate levels of craving. Once craving intensity reaches a threshold, increased craving is no longer associated with increased use. This may account for why some studies fail to observe a relationship between craving and use: craving studies often purposely focus on high levels of craving and thus are, in effect, studying the flat part of the curve in the relationship.

The question of whether cue-induced or abstinence-induced cravings are more important in maintaining appetitive behavior has not been well addressed to date. As noted earlier, it is standard practice to simply assess the level of craving irrespective of how it is induced. Findings from observational field studies suggest that cue-induced cravings may be particularly important in determining the outcome, or at least the failure, of a drug cessation attempt. The regular/timely drug dosing patterns observed when individuals are not interested in maintaining abstinence, however, suggest that abstinence-induced cravings may be more important in terms of maintaining the behavior during such periods. More work needs to be done in this area to determine whether there are functional differences between cue-induced and abstinence-induced cravings in terms of maintaining appetitive behaviors, or whether they are simply the same experience provoked via different mechanisms.

SUMMARY

Craving is an important, albeit nebulous, contributor to the maintenance of drug use and other appetitive behaviors. Its strict definition may be the subject of continued debate, but its role in driving behavior has been, and continues to be, upheld by both subjective reports and scientific reviews. Conceptually, craving can be seen as a final-common-pathway expression of motivations for such behaviors.

SEE ALSO

Cue Reactivity, Deprivation, Craving, and Affect: Intersecting Constructs in Addiction, Neural Correlates of Craving for Psychoactive Drugs, Defining and Assessing Drug Craving

List of Abbreviation

NRT nicotine replacement therapy

Glossary

Ecological momentary assessment (EMA) measurement in real time, in the natural environment.

Further Reading

Carter, B.L., Tiffany, S.T., 1999. Meta-analysis of cue-reactivity in addiction research. Addiction 94, 327–340.

Ferguson, S.G., Shiffman, S., 2009. The relevance and treatment of cue-induced cravings in tobacco dependence. Journal of Substance Abuse Treatment 36, 235–243.

Sayette, M.A., Shiffman, S., Tiffany, S.T., et al., 2000. The measurement of drug craving. Addiction 95 (Suppl. 2), S189–S210.

Tiffany, S.T., 1990. A cognitive model of drug urges and drug-use behavior: role of automatic and nonautomatic processes. Psychological Review 97, 147–168.

West, R., Schneider, N., 1987. Craving for cigarettes. British Journal of Addiction 82, 407–415.

CHAPTER

50

International Data on the Prevalence and Correlates of Comorbid Substance Use and Psychiatric Disorders

Katherine L. Mills and Christina Marel

University of New South Wales, Sydney, NSW, Australia

OUTLINE

The term comorbidity refers to the co-occurrence of two or more disorders or diseases in an individual within a defined period of time. Although the focus of this chapter is on the co-occurrence of substance use and mental disorders, it is important to note that there are many different types of comorbidity evident among individuals with substance use disorders (SUDs). Indeed, the most common form of comorbidity among people with SUDs is the occurrence of more than one SUD. Other conditions that are often found to co-occur with SUDs are physical health conditions (e.g. cirrhosis, hepatitis, heart disease, diabetes), intellectual and learning disabilities, cognitive impairment, and chronic pain.

The co-occurrence of mental disorders with alcohol, tobacco, and other drug disorders is not a new phenomenon; however, there has been an increase in awareness of this phenomenon due to the development of structured diagnostic interviews, and their use in large-scale epidemiologic surveys of mental health. Epidemiological surveys provide the opportunity to observe patterns of psychiatric comorbidity in the general population, providing valuable information about the need for services within that population. They also overcome biases inherent in studies of treatment-seeking samples.

Over the last three decades, a growing number of representative population surveys utilizing fully structured diagnostic interviews have been conducted around the world. These surveys have provided invaluable data on the prevalence and correlates of comorbid mental health and SUDs. The vast majority of research examining the population prevalence of comorbid mental health and SUDs has come from the United States, but studies have also been conducted in Australia

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00050-4

TABLE 50.1 Substance Use and Mental Health Comorbidity Surveys

Study name	Year	*N*	Age range	Response rate (%)
United States				
ECA	1980–1984	20 291	18+	76
NCS	1990–1992	8098	15–54	83
NS-R	2001–2003	9282	18+	71
NESARC (Wave 1)	2001–2002	43 093	18+	81
NESARC (Wave 2)	2004–2005	34 653	18+	87
Europe				
NEMESIS	1996	7076	18–64	70
ESEMed	2001–2003	21 425	18+	61
Australia				
NSMHWB	1997	10 641	18+	78
NSMHWB	2007	8841	16–85	60
New Zealand				
NZMHS	2003–2004	12 992	16+	74

and New Zealand, and Europe (Table 50.1). In addition to these studies, a number of population surveys examining mental health and substance use have been conducted in other countries from the America, Africa, the Eastern Mediterranean, Europe, the Western Pacific, and South east Asia as part of the World Mental Health (WMH) Survey Initiative. At the time of writing, however, only data from the aforementioned countries had been published in relation to comorbidity. This chapter presents a summary of the major findings of these studies with regard to the prevalence and correlates of comorbidity. The strengths and limitations of these studies are also discussed, as well as future directions for research.

UNITED STATES

Five large-scale epidemiologic surveys have been conducted in the United States since 1980. These include the Epidemiologic Catchment Area (ECA) survey, the National Comorbidity Survey (NCS) and its replication, as well as Waves I and II of the National Epidemiologic Study of Alcohol and Related Conditions.

Epidemiological Catchment Area Survey

The first survey designed to derive population prevalence estimates of substance use and mental disorders was the landmark ECA survey. Participants were derived from households as well as institutional settings including long-term mental hospitals, nursing homes, and penal institutions, from five catchment areas in the United States. The diagnostic interview schedule (DIS) was used to estimate the prevalence of selected Diagnostic Statistical Manual Version III (DSM-III) mood disorders (mania, dysthymia, bipolar disorder, major depression, bipolar disorder), anxiety disorders (obsessive–compulsive disorder (OCD), phobia, panic), SUDs (alcohol and drug abuse and dependence), psychotic disorders (schizophrenia and schizophreniform), somatization, antisocial personality, and anorexia nervosa.

The ECA documented the high prevalence of substance use and mental health disorders, and substantial comorbidity between them. One-third of the sample met criteria for either a substance use or mental disorder during their lifetime, and 54% of those met criteria for more than one. The odds of having a comorbid mental disorder were 2.7 times higher among those with an SUD compared to the rest of the population. Over one-third (37%) of those with an alcohol disorder, and 53% of those with a drug use disorder, also met criteria for a mental disorder during their lifetime. Conversely, 29% of those with a mental disorder met criteria for an SUD (22% alcohol, 15% other drugs).

In terms of specific SUDs, mental disorder was most prevalent among individuals with cocaine, barbiturate, and hallucinogen use disorders (76, 75, and 69%, respectively), followed by those with opiate, amphetamine, marijuana, and alcohol use disorders (65, 63, 50, and 37%, respectively). Strong associations were observed with comorbid ASPD (odds ratios (ORs) 8.3–29.2), schizophrenia (ORs 3.3–13.2), affective (ORs 1.9–6.6), and anxiety disorders (ORs 1.5–5.0).

The ECA study pioneered epidemiologic research pertaining to mental health in many ways. However, as the sample was not nationally representative, there were concerns regarding the generalizability of the study findings to the broader population. These concerns were addressed a decade later, when the NCS was conducted.

National Comorbidity Survey

The first national representative epidemiologic survey to be conducted in the United States was the NCS. As the name suggests, one of the primary aims of the NCS was to examine patterns of comorbidity. The NCS interviewed a nationally representative noninstitutionalized sample of men and women from the 48 contiguous states. A modified version of the Composite International Diagnostic Interview (CIDI) version 1.1 was used to assess for the presence of selected Diagnostic Statistic Manual Version III Revised (DSM-III-R) mood disorders (major depression, mania, dysthymia),

anxiety disorders (panic disorder, agoraphobia, social phobia, generalized anxiety disorder (GAD)), SUDs (alcohol and drug abuse and dependence), and antisocial personality disorder (ASPD) during respondents' lifetime. Nonaffective psychosis (a category made up of schizophrenia, schizophreniform disorder, schizoaffective disorder, delusional disorder, and atypical psychosis) was determined using a combination of the CIDI interview and structured clinical interviews. The past 12-month diagnoses were made in the subsample of respondents who qualified for the lifetime diagnosis and who reported at least one symptom in the 12 months prior to interview.

Like the ECA, the NCS documented high rates of comorbidity between substance use and mental health disorders. Forty-eight percent of the sample met criteria for at least one mental or SUD during their lifetime and 29% in the preceding 12 months. Just 21% of all the lifetime disorders occurred in respondents with a lifetime history of just one disorder, meaning that the vast majority of lifetime disorders in this sample (79%) were comorbid.

The strongest association observed was between substance use and antisocial personality disorder (OR 13.9). The likelihood of having a co-occurring affective or anxiety disorder in one's lifetime or in the last 12 months was 2–3 times higher among individuals with SUDs. Associations were particularly strong in relation to mania and GAD. The odds of receiving a lifetime diagnosis of nonaffective psychosis were 3.7 times higher among those with SUDs. All disorders were more strongly associated with dependence, rather than abuse. Associations also tended to be stronger in relation to drug use disorders as opposed to alcohol use disorders. Examination of the temporality of disorders revealed that among people with a history of both a mental and SUD, the mental disorder typically occurred first.

Examination of the correlates of comorbidity revealed that the likelihood of having three or more disorders within one's lifetime was positively associated with being female, of a younger age cohort, lower levels of education, and lower income. Conversely, African-American ethnicity was associated with a decreased likelihood of lifetime comorbidity. The same correlates were observed for current comorbidity, with the exception of race; current comorbidity was more likely to be present among individuals of Hispanic ethnicity. Current comorbidity was also more likely to occur among those living in major metropolitan areas.

National Comorbidity Survey-Replication

The NCS-R was conducted between 2001 and 2003, providing some of the most recent estimates of the prevalence and correlates of Diagnostic and Statistical Manual IV (DSM-IV) mental disorders in the United States. Using the WMH–CIDI, the NCS-R measured the lifetime prevalence of a wide range of mental health disorders, as well as the severity of these disorders, among a nationally representative sample of respondents from the contiguous states. The disorders included were anxiety disorders (panic, GAD, agoraphobia without panic, specific, social, post-traumatic stress disorder (PTSD), OCD, separation anxiety disorder), mood disorder (major depression, dysthymia, bipolar I or II), impulse control disorders (oppositional defiant, conduct, attention deficit/hyperactivity, intermittent explosive), and SUDs (alcohol and drug abuse and dependence). Like the NCS, past 12-month diagnoses were made in the subsample of respondents who qualified for lifetime diagnosis and who reported at least one symptom in the 12 months prior to interview.

The NCS-R estimated that 46% of the population met criteria for one or more DSM-IV disorders at some point in their lives, and 26% met criteria for a disorder in the preceding 12 months. The NCS-R found a strong association between substance dependence and mental disorders. The relationship between all mental disorders and any substance use was significant. The strongest association was with externalizing disorders (i.e. conduct disorder, oppositional defiant disorder, attention deficit-hyperactivity disorder, intermittent explosive disorder). After controlling for age, gender, and ethnicity, people with an externalizing disorder were over four times more likely to meet criteria for lifetime dependence on alcohol, nicotine, or other drugs. Individuals with mood and anxiety disorders were 2.5–3.5 times more likely to meet criteria for lifetime dependence. As in the NCS, associations tended to be stronger in relation to drug dependence as opposed to alcohol or nicotine dependence. Further, the onset of mental disorder was more likely to occur prior to the onset of dependence, particularly in relation to externalizing and anxiety disorders. Mood disorders tended to have their onset both before and after the onset of dependence.

Demographic correlates of comorbidity have received little attention in analyses of the NCS-R data to date. Similar to the NCS, however, comorbidity was positively associated with younger age cohorts. In regard to clinical correlates, the NCS has made an important contribution by examining the association between severity of illness and comorbidity. In particular, this study found that disorder severity was strongly associated with comorbidity, even though the temporally primary disorder was often relatively mild. Twenty-five percent of those with two disorders and one-half of those with three or more disorders were classified as having serious mental health disorders.

It should be noted that the version of the CIDI used in the NCS-R contained a serious fault in that respondents were required to meet criteria for abuse before being asked about dependence. This is of concern as this condition is not a requirement of the DSM-IV and a significant number of people have been shown to meet dependence criteria without meeting criteria for abuse. Thus, estimates of prevalence and associations between disorders may be biased.

National Epidemiologic Survey on Alcohol and Related Conditions

The National Epidemiological Survey on Alcohol and Related Conditions (NESARC) was conducted between 2001 and 2002 (Wave I), and a follow-up of the same respondents occurred between 2004 and 2005 (Wave II). NESARC was a nationally representative sample of those living in households or group quarters (i.e. boarding houses, non-transients hotels/motels, shelters, group homes) from all US states (contiguous and noncontiguous). The group quarters sampling frame captured important subgroups of the population with heavy substance use patterns that are often not included in general population surveys. With sample sizes of 43 093 and 34 653 in Waves I and II, respectively, the NESARC represents the largest comorbidity survey conducted to date.

The Alcohol Use Disorder and Associated Disabilities Interview Schedule-DSM-IV version (AUDADIS-IV) assessed for the presence of mood disorders (major depression, bipolar I and II, and dysthymia), anxiety disorders (panic with and without agoraphobia, social and specific phobia, GAD), and SUDs (alcohol and drug abuse and dependence). Full diagnoses were obtained for both lifetime and 12-month disorders. Personality disorders assessed on a lifetime basis at Wave I included DSM-IV avoidant, dependent, obsessive–compulsive, paranoid, schizoid, and antisocial. Additional personality disorders assessed at Wave II included borderline, schizotypal, and narcissistic personality disorders. NESARC was the first US population survey to examine the prevalence of personality disorders beyond ASPD.

Analyses of data from Wave I of the survey found that individuals with current alcohol use disorders were significantly more likely to meet criteria for a co-occurring mood, anxiety, or personality disorder, even after controlling for demographic characteristics and the presence of other comorbid psychiatric disorders (ORs 1.4, 1.3, and 1.4, respectively). After controlling for the aforementioned factors, individuals with drug use disorders were also more likely to meet criteria for a co-occurring mood (OR 1.8) or personality disorder (OR 2.2), but they were no more likely to experience an anxiety disorder. Similar ORs were observed with respect to the personality disorders assessed at Wave II. Again, comorbidity was more strongly associated with dependence than abuse, and with drugs as opposed to alcohol.

Unlike the ECA and the NCS surveys, the NESARC was able to differentiate between substance-induced and independent mood and anxiety disorders. It was estimated that 20% of individuals with an SUD met criteria for at least one independent mood disorder during the same 12-month period, and 18% met criteria for at least one independent anxiety disorder. Less than 1% of comorbid mood and anxiety disorders were classified as being substance-induced. Twenty-nine percent of those with an alcohol use disorder and 48% those with a drug use disorder met criteria for a personality disorder.

EUROPE

Netherlands Mental Health Survey and Incidence Study

The Netherlands Mental Health Survey and Incidence Study (NEMESIS) was a study of 7076 respondents from 90 Dutch municipalities in 1996. In keeping with the NCS, the NEMESIS used the CIDI version 1.1 to assess selected DSM-III-R disorders, including mood disorders (depression, dysthymia, bipolar), anxiety disorders (panic, agoraphobia, simple, social, GAD, obsessive–compulsive), SUDs (alcohol or drug abuse and dependence), eating disorders (anorexia and bulimia nervosa), and schizophrenia.

Forty-one percent of respondents met criteria for at least one mental or SUD during their lifetime: 23% in the past 12 months. Forty-five percent of those with a psychiatric disorder had suffered more than one. The overall population prevalence comorbidity between substance use and anxiety and/or affective disorders was estimated to be 2%. One-quarter of those with a current SUD had a comorbid disorder (15% mood disorder, 19% anxiety disorder). Conversely, 17% of those with a mood disorder and 13% of those with an anxiety disorder had a comorbid SUD.

The correlates of current comorbidity were similar to those observed in the NCS. The likelihood of having two or more disorders within the past 12 months was associated with being female, younger age cohorts, lower levels of education and income, being out of the workforce, and living in an urban area. Additionally, the NEMESIS found that comorbidity was more likely to occur among those whose parents had a history of psychiatric disorder and those with a history of abuse or neglect.

Further analyses were undertaken examining the correlates of specific comorbidities. Among those with

SUDs, comorbid mood and anxiety disorders were more likely to occur among females, those with lower levels of education, those whose parents had a history of psychiatric disorder, and those with a history of abuse or neglect. Comorbid mood disorders were also more likely to occur among those aged between 35 and 54 years and those who were unemployed.

European Study of the Epidemiology of Mental Disorders: Belgium, France, Germany, Italy, Netherlands, and Spain

The European Study of the Epidemiology of Mental Disorders (ESEMeD) was a cross-sectional community survey administered between 2001 and 2003 to 21 425 respondents across six European countries (Spain, Italy, Germany, Belgium, France, and Netherlands). It is the largest and most comprehensive population-based epidemiological study conducted in Europe to date. Similar to the other epidemiological studies discussed in this chapter, the ESEMeD used the WMH-CIDI to assess the prevalence of substance use and mental disorders. Current (past 12 months) comorbidity was examined in relation to selected DSM-IV mood (major depression and dysthymia), anxiety (GAD, social phobia, specific phobia, PTSD, agoraphobia, and panic disorder), and alcohol use disorders (abuse and dependence).

Twenty-three percent of those with an alcohol disorder also had a co-occurring mood or anxiety disorder, which translates into an overall population prevalence of 0.3%. Alcohol dependence was more frequently comorbid with other disorders (28%) than alcohol abuse (21%). Alcohol abuse and dependence were associated with an increased likelihood of most mood and anxiety disorders. Neither abuse nor dependence was associated with comorbid dysthymia or PTSD. Alcohol dependence was also not related to social phobia.

Demographic correlates of comorbid substance use and mental disorders were examined, however, no significant relationships were observed. This lack of difference most likely reflects a lack of statistical power to detect differences, rather than a true absence of differences.

AUSTRALIA

Australia has conducted two National Surveys of Mental Health and Wellbeing (NSMHWB). The 1997 survey examined the 12-month prevalence of selected International Classification of Diseases 10th revision (ICD-10) anxiety disorders (panic, agoraphobia, social phobia, OCD, PTSD, GAD), affective disorders (major depression, dysthymia, bipolar), and SUDs (alcohol and drug abuse/harmful use and dependence) using the CIDI version 2.1. The 2007 NSMHWB assessed for the lifetime prevalence of the same disorders using the CIDI version 3.0. Past 12-month diagnoses were made in the subsample of respondents who qualified for the lifetime diagnosis and who reported at least one symptom in the 12 months prior to interview.

In 1997, it was estimated that 23% of the population met ICD-10 criteria for a substance use or mental disorder, and approximately 25% of those met criteria for more than one class of disorder. Close to one-half (48%) of females and 34% of males with an alcohol use disorder had a comorbid affective or anxiety disorder. Two-thirds of both males and females with a drug use disorder met criteria for a comorbid affective or anxiety disorder. In terms of overall population prevalence, these figures translate to 3% of males and 2% of females meeting criteria for a comorbid substance use and affective and/or anxiety disorder within a 12-month period.

Similarly, the 2007 survey estimated the prevalence of any lifetime mental disorder to be 45% and the prevalence of any 12-month mental disorder to be 20%. One in four people (25%) with a 12-month mental disorder were diagnosed with more than one disorder. One in five (21%) of those with an SUD also met criteria for an affective disorder and one in three (33%) met criteria for an anxiety disorder. Analogous to the 1997 survey, the overall population prevalence of comorbid substance use and affective and/or anxiety disorder was estimated to be 3% for males and 2% for females.

In both surveys, females were more likely to experience comorbid anxiety and affective disorders, whereas males were more likely to experience these disorders in combination with a SUD. Further analyses of the demographic correlates of comorbidity in the 1997 survey found that having three or more disorders within the last 12 months was associated with younger age cohorts, being unmarried, and not in the workforce. Comorbidity was not associated with sex, level of education, ethnicity, or urbanicity. Analyses of the 2007 survey demonstrated a strong relationship between comorbid mental disorders and severity. More than half (54%) of those with comorbid classes of mental disorder experienced severe levels of impairment, compared to only 7% of those with one class of mental disorder.

NEW ZEALAND

New Zealand conducted its first nationally representative epidemiological study in 2003 as part of the World Mental Health Survey Initiative. The interview was based on the CIDI version 3.0, which assessed selected DSM-IV anxiety disorders (panic disorder, agoraphobia

without panic, specific phobia, social phobia, GAD, PTSD, and OCD), mood disorders (major depressive disorder, dysthymia, and bipolar disorder), SUDs (alcohol and drug abuse or dependence), and eating disorders (anorexia and bulimia). Like the NCS, the past 12-month diagnoses were made in the subsample of respondents who qualified for the lifetime diagnosis and who reported at least one symptom in the 12 months prior to interview.

Thirty-nine percent of the population met criteria for a lifetime mental health or SUD; 21% in the past 12 months. Close to one-half (49%) of those who met criteria for a disorder during their lifetime met criteria for more than one. The corresponding figure for the preceding 12 months was 37%.

Close to one-third (29%) of those with a 12-month SUD had a comorbid mood disorder; 40% had a comorbid anxiety disorder. SUDs were strongly and positively associated with all mood and anxiety disorders. In particular, strong associations were observed between drug dependence and dysthymia (OR 15.6), bipolar disorders (OR 10.6), and panic disorder (OR 10.4). Alcohol dependence was most strongly related to bipolar disorders (OR 11.8), OCD (OR 11.0), and panic disorder (OR 8.8).

There was a clear relationship between increasing numbers of disorders and case severity. Sixty percent of those with three or more disorders were classified as serious, as opposed to 12% of those with one disorder. There was also a clear association between comorbidity and suicidal behavior.

DISCUSSION

Understanding the prevalence and distribution of comorbidity is important for several reasons. Recognizing the clustering of disorders within individuals gives a greater appreciation of how and in whom the burden of psychopathology is concentrated. Estimates of current comorbidity provide important information regarding the treatment needs of the population at any given time and is essential for health service planning and delivery. An understanding of lifetime patterns of comorbidity provides opportunities for the prevention of the secondary disorders.

The data presented in this chapter demonstrate the substantial comorbidity experienced by individuals with SUDs. Direct comparisons between surveys are difficult due to differences in terms of the number and types of disorders assessed, the assessment criteria and survey instrument used, the reference timepoint, and the sampling procedures (to name a few). Despite these differences, individuals with SUDs were consistently found to be at greater risk of having comorbid mood (affective) and anxiety disorders around the world. Although measured less frequently, strong associations were also found in relation to personality disorders and psychosis.

Although the overall population prevalence of comorbidity appears to be relatively small (<3%), the burden attributed to this group is considerable, with more recent surveys consistently demonstrating strong positive associations between comorbidity, severity of illness, and disability. Thus, this seemingly small proportion of the population confers a substantial cost to the community.

The population surveys reviewed in this chapter indicate that between one-third and one-half of individuals with SUDs meet criteria for a comorbid mental disorder at some point in their lives, and 20–25% will experience these comorbid disorders in any 12-month period. Given that the majority of epidemiologic surveys have limited their focus to household populations and excluded population segments in which the prevalence of mental health and SUDs are known to be high (e.g. the homeless and those living institutions and treatment facilities), these estimates are likely to be conservative. Studies specifically targeting these populations have demonstrated that comorbidity among these groups is typically the norm rather the exception.

A further reason as to why these estimates are likely to be underestimates of the true prevalence is that, for logistical reasons, epidemiologic surveys of mental health focus only on selected mental disorders, rather than the full spectrum of possible mental disorders. Thus, many disorders are not taken into account. Indeed, the estimates presented in this chapter pertain largely to the co-occurrence of mood and anxiety disorders. It should also be borne in mind that while the focus here was on individuals who meet criteria for a diagnosis of comorbid mental health disorder, there are a large number of people with SUDs who may fall into the category of "diagnostic orphans": individuals who display a number of *symptoms* of disorders while not meeting criteria for a *diagnosis* of a disorder. Although these individuals may not meet full diagnostic criteria according to the classification systems, their symptoms may nonetheless impact significantly on their functioning and require intervention.

The fact that findings have been largely consistent across surveys conducted in different parts of the world using different methodologies is evidence of the robustness of the association between comorbid substance use and mental disorders. However, the breadth of countries in which comorbidity has been examined in nationally representative population surveys is relatively small, and restricted to nations from the developed world. It is unknown to what degree these findings can be generalized to the populations of other countries, particularly

those of the developing world. Estimates and associations may vary in countries where there are alternate health-care policies and health service delivery systems. Furthermore, differences in patterns of substance use, types of substances used, and cultural understandings of mental disorder may also lead to variations. Therefore, while studies conducted internationally may be informative to nations where local data are not available, it should be viewed with caution. Further research examining substance use and mental health comorbidity in other countries, in particular developing nations, would represent a valuable addition to the literature.

The knowledge base surrounding the association between mental disorders and particular types of SUDs could also be broadened. The predominance of research to date suggests that associations between comorbid mental and SUDs are stronger in relation to dependence as opposed to abuse, and among those with drug as opposed to alcohol use disorders. However, limited analyses have been conducted comparing the patterns and distribution of mental disorder between drug classes. Studies that do distinguish between drug type indicate that associations may be stronger in relation to drugs with a high dependence liability.

The finding that mental disorders tend to have their onset prior to SUDs highlights the importance of identification and early intervention in order to prevent the progression of primary disorders and the development of subsequent SUDs. Limited data are available with regard to the correlates of comorbidity more broadly, let alone with regard to comorbidity between substance use and mental disorders specifically. The data that are available nonetheless indicate that comorbidity is positively associated with female gender, younger age cohorts, unemployment, and lower levels of income. Findings regarding education, urbanicity, and ethnicity have been mixed. Consistent with research among clinical samples, a history of childhood abuse or neglect and parental psychiatric history appears to play an important role in the development of comorbid disorder. Further epidemiologic research identifying the correlates of comorbidity would be beneficial in informing the development of targeted interventions. Nonetheless, these findings point to the importance of interventions aimed at preventing primary disorders.

Needless to say, it is crucial that epidemiological information be as up-to-date and accurate as possible to reflect changes in nosology. An issue of concern for future epidemiologic surveys is the need for adequate response rates. As evident from Table 50.1, response rates have been declining in recent years. This phenomenon is not restricted to surveys of mental health and has been observed in relation to general population surveys more broadly. Declining response rates raise concerns regarding the validity of the estimates and associations derived, as the likelihood of response bias is increased (e.g. those with mental illness having higher survey refusal rate than those without or vice versa). It is therefore important that future epidemiologic surveys undertake measures to increase response rates so that the estimates and associations derived are representative of the population under investigation.

CONCLUSION

The epidemiologic surveys discussed in this chapter have provided crucial information about the patterns and distribution of comorbid mental health and SUDs in the United States, Europe, Australia, and New Zealand. Comorbidity between substance use and mental disorders is a common occurrence, but a relatively small group bears an especially heavy burden. Despite the frequent occurrence of comorbid mental disorders among people with SUDs, much remains to be understood about comorbidity. In particular, little is known regarding the patterns and distribution of comorbidity in countries outside of the developed world, and there is a notable lack of epidemiological research focusing explicitly on the demographic or clinical correlates of substance use and mental disorders. Further research in these domains would provide useful information for the recognition of at risk individuals, the planning of population health initiatives, and the provision of appropriate interventions.

SEE ALSO

Models of Relationships between Substance Use and Mental Disorders, Cannabis Use and the Development and Maintenance of Psychosis, Impact of Substance Use on the Course of Serious Mental Disorders, Substance Use and Mood Disorders, Substance Use in Response to Anxiety Disorders

List of Abbreviations

ASPD	antisocial personality disorder
AUDADIS-IV	The Alcohol Use Disorder and Associated Disabilities Interview Schedule Version IV
CIDI	Composite International Diagnostic Interview
DIS	diagnostic interview schedule
DSM-III	Diagnostic Statistical Manual Version III
DSM-III-R	Diagnostic Statistic Manual Version III Revised
DSM-IV	Diagnostic and Statistical Manual IV
ECA	Epidemiological Catchment Area
ESEMed	European Study of the Epidemiology of Mental Disorders
GAD	generalized anxiety disorder
ICD-10	International Classification of Diseases 10th revision

NCS	National Comorbidity Survey
NCS-R	National Comorbidity Survey-Replication
NEMESIS	Netherlands Mental Health Survey and Incidence Study
NESARC	National Epidemiologic Survey on Alcohol and Related Conditions
NSMHWB	National Surveys of Mental Health and Wellbeing
OCD	obsessive–compulsive disorder
OR	odds ratio
PTSD	post-traumatic stress disorder
SUDs	substance use disorders
WMH	World Mental Health

Further Reading

Andrews, G., Slade, T., Issakidis, C., 2002. Deconstructing current comorbidity: data from the Australian National Survey of Mental Health and Well-Being. British Journal of Psychiatry 181, 306–314.

Glantz, M.D., Anthony, P.A., Berglund, L., et al., 2009. Mental disorders as risk factors for later substance dependence: estimates of optimal prevention and treatment benefits. Psychological Medicine 39, 1365–1377.

Graaf, R., Bijl, R.V., Smit, F., et al., 2002. Risk factors for 12-month comorbidity of mood, anxiety, and substance use disorders: findings from the Netherlands Mental Health Survey and Incidence Study. American Journal of Psychiatry 159, 620–629.

Grant, B.F., Stinton, F.S., Dawson, D.A., et al., 2004a. Co-occurrence of 12-month alcohol and drug use disorders and personality disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Archives of General Psychiatry 61, 361–368.

Grant, B.F., Stinton, F.S., Dawson, D.A., et al., 2004b. Prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Archives of General Psychiatry 61, 807–816.

Hall, W., Degenhardt, L., Teesson, M., 2009. Understanding comorbidity between substance use, anxiety and affective disorders: broadening the research base. Addictive Behaviors 34, 526–530.

Jané-Llopis, E., Matytsina, I., 2006. Mental health and alcohol, drugs and tobacco: a review of comorbidity between mental disorders and the use of alcohol, tobacco and illicit drugs. Drug and Alcohol Review 25, 515–536.

Kessler, R.C., Nelson, C.B., McGonagie, K.A., et al., 1996. The epidemiology of co-occurring addictive and mental disorders: implications for prevention and service utilization. American Journal of Orthopsychiatry 66, 17–31.

Kessler, R.C., Crum, R.M., Warner, L.A., et al., 1997. Lifetime co-occurrence of DSM-III-R alcohol abuse and dependence with other psychiatric disorders in the National Comorbidity Survey. Archives of General Psychiatry 54, 313–321.

Myrick, H., Brady, K., 2003. Current review of the comorbidity of affective, anxiety and substance use disorders. Current Opinion in Psychiatry 16, 261–270.

Regier, D.A., Farmer, M.E., Rae, D.S., et al., 1990. Comorbidity of mental disorders with alcohol and other drug abuse: results from the Epidemiologic Catchment Area (ECA) Study. Journal of the American Medical Association 264, 2511–2518.

Scott, K.M., McGee, M.A., Oakley, M.A., et al., 2006. Mental disorder comorbidity in Te Rau Hinengaro: the New Zealand Mental Health Survey. Australian and New Zealand Journal of Psychiatry 40, 875–881.

Teesson, M., Proudfoot, H., 2003. Comorbid Mental Disorders and Substance Use Disorders: Epidemiology, Prevention and Treatment. Australian Government Department of Health and Ageing, Canberra.

Teesson, M., Slade, T., Mills, K., 2009. Comorbidity in Australia: findings of the 2007 National Survey of Mental Health and Wellbeing. Australian and New Zealand Journal of Psychiatry 43, 606–614.

The ESEMeD/MHEDEA 2000 Investigators, 2004. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatrica Scandinavica 109 (s420), 21–27.

CHAPTER

51

Models of Relationships between Substance Use and Mental Disorders

Douglas L. Noordsy[*], *Manish K. Mishra*[*], *and Kim T. Mueser*[$]

[*]Dartmouth Medical School, Lebanon, NH, USA [$]Dartmouth Psychiatric Research Center, Concord, NH, USA

OUTLINE

Numerous studies have demonstrated an elevated prevalence of substance use disorders (SUDs) among individuals with psychiatric disorders. Strikingly, the odds of co-occurring disorders are elevated by 1.5–2 times the general population rate among individuals with depression and anxiety disorders in these samples, but the odds are elevated 3–6-fold for individuals with schizophrenia or bipolar disorder. This has motivated attempts to understand why such high risks for SUDs occur in this population.

Many explanations have been proposed for the increased risk of SUDs in people with serious mental illness. Four general models of increased comorbidity include secondary psychiatric disorder models, secondary SUD models, common factor models, and bidirectional models. According to secondary psychiatric disorder models, substance use contributes to the onset of mental illness in individuals who would otherwise not have developed these disorders. Secondary SUD models propose that serious mental illness increases the person's chances of developing a SUD. Different models may account for comorbidity in individual patients, and more than one model may apply for a given individual over time, or in relation to different substances.

Before reviewing these models we note that much of the data on the prevalence of SUDs in serious mental illness are drawn from clinical samples, including some of the data from the Epidemiologic Catchment Area study. According to Berkson's fallacy, estimates of comorbidity are inflated when samples are obtained from treatment settings, as opposed to the general population, because either disorder increases the likelihood that individuals will receive treatment. As a result, estimates of comorbidity may be inflated by sampling bias. However, it is unlikely that this factor is sufficient to explain the high rates observed.

SECONDARY PSYCHIATRIC DISORDER MODELS

The theory that substance abuse can lead to serious mental illness has focused primarily on stimulants, hallucinogens, and cannabis because of their

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00051-6

psychotomimetic effects. In contrast, there is less evidence that alcohol use can contribute to onset of schizophrenia or bipolar disorder, and some debate as to whether it can mask their onset.

Models proposing that psychotomimetic drug use can lead to long-term psychotic disorders build on the catecholamine hypothesis of schizophrenia or mood disorders. Animal research has shown that repeated or continuous stimulant administration can lead to increased sensitivity of response, or behavioral sensitization. Similarly, increased electrophysiological and behavioral responses can be induced by repeated electrical stimulation or stimulant administration, referred to as kindling. Behavioral sensitization and kindling due to substance use have been suggested as mechanisms by which drug use may precipitate serious mental illness.

McLellan et al. followed 51 male veterans with at least yearly hospitalizations for drug abuse over a 6-year period. Although there were no differences in psychiatric symptoms initially, by endpoint 5 of 11 stimulant users had developed psychosis, and 8 of 14 depressant users had developed severe depression. However, if substance type predicted psychiatric diagnosis, then different diagnostic groups would tend to abuse different types of substances, which is not the case. There is no consistent association between psychiatric diagnosis and the specific types of substances abused by people with serious mental illness. Indeed, polysubstance misuse is the most common pattern.

Several studies have compared people who develop long-term serious mental illness following drug abuse to matched controls without SUDs. Most of these researches have failed to find consistent differences between these groups. Furthermore, there are no consistent differences between people with co-occurring disorders whose SUD developed before versus after the mental illness.

CANNABIS AS A PRECIPITANT OF PSYCHOSIS

Interest in the potential role of cannabis as a precipitant of psychosis was triggered by the intriguing findings of a landmark study by Andréasson and colleagues. Andréasson et al. reported that in a large 15-year prospective follow-up study of young men conscripted into the Swedish army, there was a strong association between history of cannabis use at conscription and later diagnosis of schizophrenia. Although this association was reduced when other variables were statistically controlled, it nevertheless remained significant. No associations were found between use of other drugs and the subsequent development of schizophrenia. Additional analyses indicated that the cannabis users who developed schizophrenia had a more rapid onset of illness characterized by positive symptoms than nonusers who developed schizophrenia, which the authors interpreted as supporting an etiologically distinct subgroup.

Since this study, several other population-based studies have demonstrated a relationship between cannabis use and the development of schizophrenia, controlling for possible confounders. The relationship between cannabis use and schizophrenia is dose-dependent, is stronger for earlier cannabis use, and is not modified by other drug use. Furthermore, cannabis use is associated with an earlier age of onset of psychosis in this population. In addition, Caspi and colleagues reported that adolescent cannabis use elevated risk for subsequent schizophreniform disorder, with use by age 15 conveying even higher risk than use by age 18, but only among those with a genetic marker for schizophrenia risk (see Section Genetic Factors). These studies are consistent with the stress–vulnerability model of schizophrenia, in which substance use may trigger the psychotic disorder in people who are vulnerable because of genetic or neurodevelopmental risk, where substance abuse serves as just one of many potential influences.

Neurobiological hypotheses were proposed by Müller-Vahl and Emrich to explain the relationship between cannabinoids and schizophrenia. The authors suggested that cannabis could influence the onset of psychosis through either exogenous or endogenous mechanisms. The exogenous hypothesis proposed that cannabis abuse could alter the endocannabinoid system, resulting in the onset or exacerbation of symptoms associated with schizophrenia. The endogenous hypothesis proposed that a disrupted endogenous cannabinoid system could influence the degree of psychosis experienced by patients with schizophrenia. This would support the notion that cannabis is a risk factor for either new onset of psychosis or exacerbation of symptoms in genetically predisposed people. In both of these mechanisms, the dysregulation of the cannabinoid system has potential to modify or interact with other neurotransmitter systems. Evolving understanding of the role and functions of the endocannabinoid system may help to reconcile the "Cannabinoid Hypothesis" described above with the well-described glutamatergic and dopaminergic systems that are associated with schizophrenia.

A central question that this research raises is whether cannabis use precipitates an earlier age of schizophrenia onset in people who would have developed the illness at a later age, or whether it causes schizophrenia in some vulnerable people who would otherwise not have developed the illness. If cannabis use causes schizophrenia in some people, one might expect to see increase in the prevalence of schizophrenia in countries where cannabis

use has increased. Two studies that addressed this question in birth cohorts in Australia between 1940 and 1979, and in the United Kingdom from 1996 to 2005, failed to find such an association. However, questions about the reliability of the diagnosis of schizophrenia over the time periods studied means that a possible increase in the incidence of schizophrenia due to the growth of cannabis use in a population cannot be ruled out.

Hickman and colleagues took a different approach to exploring the public policy implications of the association between cannabis use and the onset of psychosis. These researchers addressed the question of how many cannabis users would need to be prevented in order to prevent a single case of schizophrenia or psychosis, based on data collected in England and Wales. The results indicated that for men between 1007 and 8416 (depending on age) heavy cannabis users would need to be prevented, and for women between 1408 and 22 732 heavy cannabis users would need to be prevented, to prevent one case of schizophrenia. Approximately four to five times more cases of light cannabis use would need to be prevented to prevent the onset of schizophrenia. The results of this study suggest that the public health importance of preventing cannabis use to prevent schizophrenia is not clear.

Although the role of cannabis use in the precipitation of psychosis is unclear, there is strong evidence that cannabis interacts with the course of the illness. A large study based in Suffolk County, New York, suggested that there was an association between cannabis use and an adverse course of psychotic symptoms in patients who were newly diagnosed with schizophrenia spectrum disorders. This study followed psychotic symptoms over a 10-year follow-up period and is the longest assessment of its kind to date. It included over 200 participants recruited from 12 regional inpatient facilities. There were several findings in this work that provide insight into the relationship between cannabis and positive symptoms of schizophrenia. Of note, a positive relationship was established between the severity of psychotic symptoms and cannabis use, consistent with other studies. The mathematical modeling in the Foti study revealed a bidirectional nature in predicting outcomes. That is to say, not only did patients with prior cannabis use have more severe psychotic symptoms at the next assessment interval, but more severe psychotic symptoms also predicted future use of cannabis. A further implication is that there may be a reduction in the severity of psychotic symptoms with successful cessation of cannabis use.

While support for a simple causal role for cannabis use in schizophrenia is inconclusive, evidence is stronger that substance use can precipitate schizophrenia in vulnerable individuals. Both people with schizophrenia that was preceded by SUD and those with only schizophrenia have stronger family histories of schizophrenia than people with only an SUD. In addition, drug abuse not specific to cannabis is associated with an earlier age of schizophrenia onset.

SUBSTANCE USE AND PRECIPITATION OF BIPOLAR DISORDER

Of the several comorbidities that have been associated with substance abuse, bipolar disorder has been shown to have the highest prevalence as a concurrent diagnosis. Swann and colleagues suggested that individuals who have co-occurring bipolar and SUDs tend to have a more severe course of bipolar disorder, defined as having an earlier onset of symptoms, having more frequent episodes, and resulting in more complications. The complications included more anxiety, aggressiveness, legal involvement, and suicidal behavior. The search for a shared mechanism to explain the correlation between the two diagnoses identified many common characteristics including impulsivity, susceptibility to behavioral modification, and poor modulation of motivation and response to stimulus.

Another study by Kemp et al. demonstrated an association between rapid cycling bipolar patients with a comorbid substance use disorder and poor treatment response to mood stabilizers. Having a substance abuse comorbidity also increased the likelihood of incurring serious medical problems. The ability to predict associated risks and therapeutic outcomes may guide treatment options for this increasingly common population of people with co-occurring disorders.

The strong evidence that substance use worsens the course of bipolar disorder raises the question of whether it can also precipitate onset of the illness. People whose alcohol use disorder developed before their bipolar disorder have been found to have a later age of onset of bipolar disorder than those whose alcoholism came second. Furthermore, lower familial rates of bipolar disorder have been found in people whose alcoholism antedated their bipolar disorder, as well as fewer affective episodes and a more rapid recovery, compared with those whose bipolar disorder came first. These findings suggest that alcohol misuse may precipitate first episodes of mania in some persons who might not otherwise develop the disorder, or may have developed it at a later age.

SECONDARY SUBSTANCE USE DISORDER MODELS

A variety of different models posit that serious mental illness increases individuals' vulnerability to developing

an SUD, including alleviation of symptoms/dysphoria, multiple risk factors, supersensitivity, and brain reward circuit dysfunction.

Alleviation of Symptoms/Dysphoria Model

The term "self-medication" to refer to people's use of substances to alleviate psychiatric symptoms or other negative mood states has been used in so many different ways that there is little consensus in the scientific community as to its meaning. To avoid confusion, we use the term "alleviation of symptoms/dysphoria model" for this broad set of explanations for high rates of SUDs in people with serious mental illness. One version of this model suggests that substances are selected by the individual based on their specific effects on symptoms. Although research indicates that people are aware of their initial psychological reactions to substances (Mueser et al., 1995), there is little evidence for substance selection related to specific diagnoses or internal states. Self-report studies find that people with co-occurring disorders report that alcohol and other substances alleviate social problems, insomnia, depression, and other problems across diagnoses, but rarely report that specific substances alleviate specific symptoms. Studies of clinical epidemiology show that individuals with serious mental illness use the same substances as others in society but at higher rates; substance selection is not related to diagnosis but is related to availability and market forces.

Another more general version of this model is that people with serious mental illness are prone to using substances in response to distress. In support of this, research shows that people with serious mental illness often experience dysphoria, and self-reports of people with co-occurring disorders indicate that the alleviation of unpleasant feelings is a common motive for using substances. Furthermore, there is some evidence that people with co-occurring disorders have higher levels of dysphoria than those with serious mental illness alone.

Multiple Risk Factors Model

Possible indirect mechanisms involve a variety of risk factors related to serious mental illness, such as social isolation, poor interpersonal skills, poor cognitive skills, school and vocational failure, poverty, lack of adult role responsibilities, lack of structured daily activities, association with substance-using subgroups, and living in neighborhoods with high rates of drug availability. This model is consistent with Kandel's work demonstrating that achievement of adult social roles and responsibilities was associated with discontinuation of substance use among young adults (role incompatibility theory). Little research has addressed multiple risk factor models, but self-reports regarding reasons for use are consistent with the identified factors.

Supersensitivity Model

This model is an elaboration of the stress–vulnerability model for schizophrenia, which proposes that environmental stress interacts with psychobiological vulnerability to precipitate the onset of serious mental illness or to trigger relapses. Because vulnerability is defined in terms of increased biological sensitivity to stress, it may also apply to the effects of alcohol and drugs. This sensitivity may render people with serious mental illness more likely to experience negative consequences from using relatively small amounts of substances.

A number of avenues of research provide support for this model. First, people with co-occurring disorders tend to misuse lower quantities of substances than those with primary SUDs, and are less likely than primary SUD populations to develop physical dependence. Second, in pharmacological "challenge" tests, people with serious mental illness are highly sensitive to low doses of psychoactive substances that produce minimal responses in controls. Third, people with serious mental illness are sensitive to negative clinical effects such as symptom relapses following use of small quantities of alcohol or drugs. Fourth, Drake and Wallach reported that fewer than five percent of people with serious mental illness were able to sustain symptom-free drinking over time without negative consequences, in marked contrast to approximately fifty percent of the general population who drink alcohol over time without developing a disorder, suggesting increased sensitivity to the effects of alcohol.

Brain Reward Circuit Dysfunction Model

A number of authors have noted overlap in the neural circuitry putatively involved in both SUD and schizophrenia. This has led to a model that theorizes that biological vulnerability to SUD is inherent to the neurobiology of schizophrenia. Substances of abuse activate dopaminergic mesocorticolimbic tracts involved in reward, and repeated use increases sensitivity to substance effects. The reward circuit regulates repetition of behavior. People with schizophrenia appear to have a dysregulation of this circuit, which results in blunted responses to rewarding stimuli. The hypothesis suggests that people with schizophrenia are vulnerable to SUDs because substances of abuse directly stimulate this circuit, creating relatively greater reward response and stimulating compulsive repetition of the behavior of substance use. Similar reward circuitry dysfunction hypotheses have been put forward to explain the

etiology of SUD in other contexts. The reward dysfunction argument is consistent with research showing that substances with abuse potential increase dopaminergic turnover in reward circuitry in both animals and humans, as well as data indicating abnormal reward circuitry responsiveness to drug cues in people with primary SUDs.

Several areas of research have provided indirect support for the hypothesis in people with schizophrenia. For example, people with schizophrenia do not show the normal increase in P300 event-related potential amplitude to stimuli associated with monetary reward, have abnormal hedonic and brain activation responses to odors, and have structural and functional brain abnormalities in their reward circuitry, including frontal lobe subregions, striatum, amygdala, and hippocampus. This model has been primarily been applied to schizophrenia, but may have relevance in bipolar disorder as well, to the degree that it shares neurobiologic underpinnings with schizophrenia.

COMMON FACTOR MODELS

Common factor models posit that high rates of comorbidity are the result of shared vulnerabilities to both disorders. To the extent that specific factors can independently increase the risk of developing both disorders, increased comorbidity can be explained. Two risk factors have been studied in some detail: genetics and antisocial personality disorder. Other factors have been identified but less thoroughly evaluated as reasons for increased rates of comorbidity, including socioeconomic status and cognitive impairment.

Genetic Factors

Family history and twin studies provide strong evidence that genetic factors contribute to the development of schizophrenia, bipolar disorder, and SUD, although single, common genetic causes of these disorders are unknown. The question is whether genetic vulnerability to one disorder also increases risk for another disorder. Research shows that people with comorbid disorders are more likely to have relatives with an SUD than similar people with only serious mental illness. These findings suggest that genetic vulnerability to SUDs plays a role in the development of some cases of comorbid substance use and serious mental disorders.

However, do such genetic factors account for increased comorbidity via susceptibility to both disorders? Research examining the rate of SUDs in the relatives of people with serious mental illness and the rate of serious mental illness in relatives of people with an SUD addresses this question. If shared genetic vulnerability to both serious mental illness and SUD accounted for increased comorbidity, higher rates of the other disorder would be expected in the relatives of persons with one of the disorders. There are at least two possible sources of shared genetic vulnerability in families. First, within individual family members, genetic vulnerability to one disorder could also be associated with increased vulnerability to the other disorder, with offspring at increased risk to developing both disorders from the genetic contribution of that individual parent. Second, if family members with one disorder were more likely to mate with individuals with the other disorder than would be expected by chance alone (i.e. cross-trait assortative mating), the offspring would be at increased risk to developing comorbid disorders, due to the genetic contributions from each parent.

Some research provides evidence against a simple genetic model. Several studies indicate that genetic risk to schizophrenia or bipolar disorder is not associated with an increased risk of SUD in relatives, or vice versa. However, one recent investigation of bipolar and SUD comorbidity described an association of serotonin and dopamine gene polymorphisms. This linkage study used a genomic database with 278 bipolar patients to evaluate candidate genes in dopamine and serotonin pathways that are associated with substance use and bipolar disorders. Seven dopamine receptor and serotonin receptor transporter genes were targeted for closer scrutiny. After statistical analysis was completed on the genetic data of the bipolar patients, a potential association between the 5HT2C serotonin transporter gene and substance abuse comorbidity was identified.

More complex genetic models of the relationship between co-occurring schizophrenia and SUD have been proposed. For example, Caspi et al. reported that a genetic variation in the COMT gene moderated the effect of adolescent cannabis use on the later development of psychotic symptoms or schizophrenia. Adult cannabis use did not have the same effect, highlighting that the effect of substance use on the developing brain may be different and more deleterious than the effect of substances on adult brains. This research also exemplifies that genetic risk for co-occurring disorders may be enacted via gene–environment interactions, whereby substance abuse acts as an environmental stressor on the developing brain.

Genetic susceptibility for these disorders accumulates across a multitude of genes regulating various aspects of brain function. Another line of preliminary research has examined genetic variations of the dopamine D3 receptor (DRD3) and its moderators. D3 receptors are expressed in the mesocorticolimbic dopaminergic system and are increased in postmortem studies of individuals with schizophrenia. The DRD3 receptor is highly expressed

in the nucleus accumbens, where reward, including substance-induced reward, is mediated. The DRD3 receptor modulates dopamine movement in this structure when substances are used. Although studies have been mixed, some research has shown that homozygosity in a polymorphism in the DRD3 was associated with co-occurring addiction in schizophrenia.

Antisocial Personality Disorder

Another possible common factor is antisocial personality disorder. Extensive research has shown that antisocial personality disorder and its childhood precursor conduct disorder, are strongly related to increased vulnerability to SUD, and a more severe course of addiction. In addition, similar associations have been reported between antisocial personality disorder and serious mental illness. Specifically, symptoms of conduct disorder in childhood, such as repeated fighting, truancy, and lying, have been found to be predictive of the later development of schizophrenia, and to a lesser extent, bipolar disorder. Furthermore, increased rates of antisocial personality disorder have been reported in both schizophrenia and bipolar disorder.

The strong associations between conduct disorder, antisocial personality disorder, and SUDs, and the increased prevalence of antisocial personality disorder in persons with serious mental illness, suggest a role for antisocial personality disorder as a common factor underlying increased comorbidity. More direct support is provided by evidence of people with serious mental illness and past conduct disorder or antisocial personality disorder are more likely to have comorbid SUD than similar people without antisocial personality disorder. Finally, among persons with co-occurring disorders, the additional diagnosis of conduct disorder or antisocial personality disorder is associated with a more severe course of SUD and a stronger family history of SUD, consistent with research on antisocial personality disorder in persons with primary SUD.

Thus, moderately strong evidence suggests that antisocial personality disorder is a common factor that may contribute to the increased rate of SUDs in a subset of people with serious mental illness. Further work is needed to evaluate the role of temperament and to rule out other common factors related to antisocial personality disorder that could account for its relationships with SUD and serious mental illness.

BIDIRECTIONAL MODELS

Bidirectional models suggest that ongoing interactions between serious mental illness and SUD account for increased rates of comorbidity. For example, substance abuse could trigger serious mental illness in biologically vulnerable individuals, which is subsequently maintained by continued substance use due to socially learned cognitive factors, such as beliefs, expectancies, and motives for substance use. Consistent with this model, there is evidence that SUD worsens the course of serious mental illness, and that worsening symptoms are related to heavier substance use. Despite the intuitive appeal of bidirectional models, limited research has evaluated whether bidirectional interactions lead to greater psychiatric SUD comorbidity.

CONCLUSIONS

This chapter has reviewed the evidence supporting various models that attempt to explain the elevated prevalence of SUDs among individuals with psychiatric disorders. Much work remains to be done to clarify, refine, and test these hypotheses to develop a clear understanding of the psychoneurobiology of co-occurring disorders. We conclude with an important caveat and an agenda for future research. The secondary psychiatric disorder and secondary SUD models reviewed herein are focused on understanding the high rates of comorbidity, but do not imply that treatment of the "primary" disorder is adequate care. Clinical studies consistently find that treatment approaches that deliver care for both the psychiatric and the SUDs in an integrated fashion, treating both disorders as primary with simultaneous interventions designed to minimize risks to exacerbate the other disorder, achieve superior treatment outcomes. The notion of a primary–secondary relationship between psychiatric disorders and SUDs (defined in treatment studies by the chronologic order of onset in an individual's life), in which one disorder would resolve once the other is adequately treated, is not supported by available evidence, at least when each disorder is well established.

Recent advances in the understanding of neuroplasticity and epigenetics may inspire innovation in future research into the disparity between these models of comorbidity and related clinical treatment trials. We now have access to concepts and tools that have the potential to allow an increasingly sophisticated inquiry into the relationships between psychiatric and SUDs. Such inquiry might address the challenge of understanding how a disorder that may appear to have its origins in response to another disorder of earlier onset in an individual's life course can evolve into a fully established, independent disorder that requires specific, primary treatment.

List of Abbreviations

DRD3 dopamine D3 receptor
SUD substance use disorder

Glossary

Berkson's fallacy the observation that estimates the comorbidity of any two disorders are inflated when samples are obtained from treatment settings (such as a hospital or outpatient program) than from the general population (such as a household survey) because either disorder increases the likelihood that individuals will receive treatment.

Comorbidity the co-occurrence of two or more disorders, such as psychiatric and substance use disorders.

Supersensitivity hypothesis a hypothesis that some of the high comorbidity between serious psychiatric disorders and substance use disorders is due to the increased biological sensitivity to the effects of substances on people with psychiatric disorders.

Further Reading

Andréasson, S., Allebeck, P., Engström, A., Rydberg, U., 1987. Cannabis and schizophrenia: a longitudinal study of Swedish conscripts. Lancet 2, 1483–1486.

Berkson, J., 1949. Limitations of the application of four-fold tables to hospital data. Biological Bulletin 2, 47–53.

Caspi, A., Moffitt, T.E., Cannon, M., McClay, J., Murray, R., Harrington, H., Taylor, A., Arseneault, L., Williams, B., Braithwaite, A., Poulton, R., Craig, I.W., 2005. Moderation of the effect of adolescent-onset cannabis use on adult psychosis by a functional polymorphism in the catechol-O-methyltransferase gene: longitudinal evidence of a gene X environment interaction. Biological Psychiatry 57, 1117–1127.

Chambers, R.A., Krystal, J.H., Self, D.W., 2001. A neurobiological basis for substance abuse comorbidity in schizophrenia. Biological Psychiatry 50, 71–83.

Drake, R.E., Wallach, M.A., 1993. Moderate drinking among people with severe mental illness. Hospital and Community Psychiatry 44, 780–782.

Foti, D., Kotov, R., Guey, L., Bromet, E., 2010. Cannabis use and the course of schizophrenia: 10 year follow-up after first hospitalization. American Journal of Psychiatry 167, 987–993.

Hickman, M., Vickerman, P., Macleod, J., Lewis, G., Zammit, S., Kirkbride, J., Jones, P., 2009. If cannabis caused schizophrenia – how many cannabis users may need to be prevented in order to prevent one cause of schizophrenia? England and Wales calculations. Addiction 104, 1856–1861.

Kandel, D.B., Raveis, V.H., 1989. Cessation of illicit drug use in young adulthood. Archives of General Psychiatry 46, 109–116.

Kemp, D., Gao, K., Ganocy, S., Caldes, E., Feldman, K., Chan, P., Conroy, C., Bilali, S., Findling, R., Calabrese, J., 2009. Medical and substance abuse comorbidity in bipolar disorder. Journal of Affective Disorders 116, 64–69.

Kessler, R.C., Crum, R.M., Warner, L.A., Nelson, C.B., Schulenberg, J., Anthony, J.C., 1997. Lifetime co-occurrence of DSM-III-R alcohol abuse and dependence with other psychiatric disorders in the National Comorbidity Survey. Archives of General Psychiatry 54, 313–321.

Mclellan, T.A., Woody, G.E., O'Brien, C.P., 1979. Development of psychiatric illness in drug abusers: possible role of drug preference. New England Journal of Medicine 301, 1310–1314.

Mueser, K.T., Noordsy, D.L., Drake, R.E., Fox, L., 2003. Integrated Treatment for Dual Disorders: A Guide to Effective Practice. Guilford Press, New York.

Müller-Vahl, K.R., Emrich, H.M., 2008. Cannabis and schizophrenia: towards a cannabinoid hypothesis of schizophrenia. Expert Reviews of Neurotherapeutics 8, 1037–1048.

Regier, D.A., Farmer, M.E., Rae, D.S., Locke, B.Z., Keith, S.J., Judd, L.L., Goodwin, F.K., 1990. Comorbidity of mental disorders with alcohol and other drug abuse: results from the Epidemiologic Catchment Area (ECA) study. Journal of the American Medical Association 264, 2511–2518.

Swann, A., 2010. The strong relationship between bipolar and substance-use disorder. Annals of the New York Academy of Sciences 1187, 276–293.

CHAPTER

52

Substance Use and Mood Disorders

Frances Kay-Lambkin[*,$], *Amanda Baker*[$]

[*]National Drug and Alcohol Research Center, University of NSW, Sydney, NSW, Australia

[$]Center for Brain and Mental Health Research, University of Newcastle, Callaghan, NSW, Australia

OUTLINE

In 2007, the second National Survey of Mental Health and Wellbeing (NSMHWB) was conducted in Australia, and confirmed the results of similar population-based epidemiological studies in the United Kingdom and the United States. The lifetime rates of mental disorders was 46%, 20% experienced a mental disorder in the past 12 months, and 1 in 10 met criteria for a current mental disorder, within the month prior to interview. The three most common mental disorders were anxiety, mood, and alcohol/other drug (AOD) use disorders, occurring at rates of 14, 6, and 5%, respectively, in the previous 12-month period. Over one-fifth (22%) of survey respondents identified as current tobacco smokers.

Mounting pressure is exerted on the health system by the increasing prevalence of mood and substance use disorders (SUDs). For example, depression and alcohol use disorders were ranked 3 and 17 in contribution to the global disease burden in 2008, with depression elevated to first place and alcohol use to fifth among middle-high income countries. These disorders frequently co-occur and together are associated with significant disability, premature death and morbidity on a global scale.

THE PREVALENCE OF ALCOHOL, TOBACCO, AND OTHER DRUG USE PROBLEMS

Consistently, population surveys in Western countries reveal that the most commonly used substance among adults is tobacco, followed by alcohol, marijuana, and psychostimulants (such as cocaine and methamphetamines). The focus of this chapter is alcohol and illicit drugs: issues related to tobacco will be only briefly covered.

Typically, epidemiological data exploring the prevalence of AOD use problems focuses on the formal AOD use disorders of substance abuse and/or substance dependence. For example, the US Epidemiological Catchment Area (ECA) study reported lifetime prevalence rates of alcohol abuse/dependence at around 14%, with 6% lifetime

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00052-8

prevalence of other drug abuse/dependence. Australian data indicate that one in 15 adults meet criteria for an alcohol use disorder in a 12-month period, and one in 45 adults report other drug use disorders over the same period. Marijuana was the most commonly used illicit drug in both the ECA and NSMHWB surveys, and in Australia accounted for 2% of the other drug use disorders reported by the sample. The NSMHWB attempted to estimate the prevalence of harmful/hazardous AOD use, with data indicating that in the past 12 months, 3% of respondents reported using alcohol at harmful levels, 0.6% reported harmful cannabis use, and 0.9% reported harmful use of any drug (excluding alcohol).

AOD use disorders are more common among males than females. For example, the NSMHWB revealed that the rates of alcohol dependence in the past 12 months were more prevalent in males (2% versus 1%), and rates of other drug use disorders were reported among 1% of males and 0.4% of female respondents. Despite this, rates of help-seeking for AOD use disorders are generally higher among females than males, with around 39% of women in the NSMHWB seeking help for alcohol-related problems, compared with 23% of men. The pattern of smoking tobacco that is most harmful to health is daily smoking, and the great majority of smokers do so every day. In Australia, as with AOD, males are more likely to be daily smokers (18%) than females (15%).

Aside from gender, the prevalence of AOD use disorders is affected by various other sociodemographic variables such as employment, marital status, and age. To use the NSMHWB as an example, being single and unemployed were factors associated with higher rates of alcohol use disorders, as was decreasing age of survey respondents. The same patterns were found for other drug use problems. Smoking has been found to be more common among people who are unemployed, living in areas with the least socioeconomic resources or living in remote areas. Indigenous Australians are more likely to smoke than other Australians (34 and 19% respectively). These differences are also observed in epidemiological studies internationally. For example, the National Epidemiological Survey on Alcohol and Related Conditions (NESARC) conducted in the United States revealed that younger people (aged 18–29) were 13 times more likely to report an alcohol use disorder in the 12 months prior to survey than were their older counterparts, and that single respondents and those who were currently widowed, divorced or separated were around two times more likely to report a 12-month alcohol use disorder than were their married/cohabiting equivalents. This survey also revealed that Native Americans were at increased risk of lifetime alcohol dependence, who reported double the risk of Black, Asian, and Hispanic survey respondents.

THE PREVALENCE OF MOOD DISORDERS

Mood, or affective, disorders encompass a range of symptoms and diagnoses where the underlying feature is a disruption to the person's general mood and emotions. Two groups of mood disorders are commonly recognised, depressive disorders and bipolar disorders, which are distinguished from each other by the presence of manic episode(s). A depressive episode is characterized by depressed mood and/or loss of interest or pleasure in life activities for at least 2 weeks, during that time the person also experiences at least five other key depressive symptoms. A manic episode is a period of abnormally and persistently elevated, expansive, or irritable mood lasting for at least 1 week, and accompanied by three or more key symptoms. Whilst depressive or manic episodes are not codeable disorders on their own, a major depressive disorder is diagnosed when a person experiences a single or recurrent (two or more) depressive episode in the absence of a manic episode. If periods of mania have been experienced, then a bipolar disorder is likely to be present. Bipolar disorders are further categorized according to whether the dominant feature of the illness is mania (or daily cycling between manic and depressive episodes, bipolar I disorder) or recurrent depressive episodes (with hypomanic episodes, bipolar II disorder).

Within each of these mood disorder categories, people can experience subthreshold symptoms that do not meet formal diagnostic criteria, but continue to cause significant impairments in functioning and quality of life for the individual. For example, chronic depressive disorder (dysthymia) refers to the presence of a depressed mood for a period of 2 years or more, along with two of the symptoms of a depressive episode during that time. Similarly, cyclothymic disorder is a state of cycling between hypomanic (where manic symptoms are observed by others, but symptoms are not as severe) and depressive episodes that do not meet the criteria for a diagnosis of bipolar disorder. Importantly, both depressive and bipolar disorders may be classified as substance-induced; that is, a disturbance that meets full criteria for the relevant mood disorder. However the symptoms occur or develop within a month following severe intoxication or withdrawal from alcohol/other drugs, and do not persist for more than a month following the cessation of the substance.

Epidemiological research suggests that, aside from AOD use disorders, affective disorders (namely depression) are among the most common mental disorders experienced in the general population. For example,

population studies conducted in the United States indicate that the lifetime prevalence of mood disorders is 21%.

In Australia, the NSMHWB indicated that mood disorders occurred in 6% of the study population in the previous 12 months, supporting previous population-based surveys of mental disorders conducted internationally. Depressive episodes were reported by 4% of participants in the previous 12 months, and bipolar disorders by 2% of NSMHWB respondents. Internationally, bipolar I disorders generally occur in 0.8% of the general population, and 1.1% meet criteria for bipolar II disorder in a given 12-month period.

In general, depressive disorders are more common in females than males, and tend to decrease with age, declining markedly after 55 years of age. There are also gender differences in the experience and prevalence of bipolar disorders. The onset of bipolar disorder is often later in women than men, and women tend to experience more depressive episodes and rapid cycling than do men, with bipolar II disorder appearing more commonly in females. In addition, unemployed people report higher rates of mood disorders and rates are higher if people do not complete secondary school, are separated or divorced and if they currently live alone. For example, in the NSMHWB, rates of depression rose to 11% for those people without housemates in the year prior to the survey. These patterns were not observed among people with anxiety disorders.

Health service utilisation is typically low for people with mood disorders. Because of the gender differences in the presentation of bipolar disorders, women tend to experience a delay in the diagnosis of bipolar disorder and hence in gaining access to appropriate treatment. Depression accounts for the most disability of the high prevalence disorders, with respondents to the NSMHWB indicating on average that for almost 12 days of the past 4 weeks, they were completely unable to carry out their usual activities owing to their symptoms. Further, of the 13% of disability-adjusted life years attributable to mental disorder, mood disorders alone accounted for 33%, and depression was responsible for 12% of the total burden of nonfatal global disease. Bipolar disorder accounts for more disability-adjusted life years than all forms of cancer, epilepsy, and Alzheimer's disease. This is despite effective psychological and pharmacological treatments being available for the mood disorders.

THE PREVALENCE OF COMORBID ALCOHOL/OTHER DRUG USE AND MOOD DISORDERS

As noted in other chapters in this volume, a reliable finding across population-based surveys is the common co-occurrence of mental disorders, referred to as "comorbidity." The presence of a mental disorder seems to increase the risk of developing additional mental disorders. For example, the Australian NSMHWB reported that one-quarter of people with a mental disorder experienced more than one class of mental disorder. Comorbid conditions may meet criteria for diagnostic-level disorders, according to standard clinical classification systems. However symptoms do not necessarily need to meet standard diagnostic criteria in order for comorbidity to be present, and for this comorbidity to impact significantly on a person's functioning across many domains. However, epidemiological data tend to report on disorder-level comorbidity. For example, based on data from the US National Comorbidity Survey, males who smoked cannabis were 1.6 times more likely to develop a major depressive episode than were nonusers. Similarly, females were at 2.6 times the risk of a major depressive episode if they smoked cannabis, even at nondependent levels. In research conducted among 2000 Australians aged 14–21 years, the association between marijuana use and symptoms of depression and anxiety was examined. Young adult females who used cannabis daily, had five-fold higher rates of depressive and anxiety symptoms, after controlling for the effects of other substance use. Weekly cannabis use as a teenager predicted a two-fold higher rate of depression and anxiety in females in young adulthood after controlling for baseline mental health and other potential confounders. By contrast, depression and anxiety symptoms in adolescence did not significantly predict later cannabis use.

The ECA study revealed that the rate of alcohol use disorders among people with mental disorders was 22%, compared to 11% of their counterparts without mental disorders. Rates of other drug use disorders among people with mental disorders were around 15%, relative to a rate of 4% of the remaining survey respondents. Taken together, these findings suggest that adults with lifetime mental disorders are twice as likely to experience alcohol use disorders and around four times more likely to have a drug use disorder than are their counterparts without mental disorders.

Equally, studies have indicated an increased rate of psychiatric disorder is likely among people using AOD. Although rates of some acute mental disorders among AOD users are comparable to the general population, a consistent finding has been that rates of depressive disorders are elevated among AOD users, particularly those with alcohol abuse and dependence. For example, in the US NESARC, 20% of people with an AOD use disorder had a co-occurring mood disorder, and 20% of people with mood disorder also had a SUD. Major depression occurred in 10% of national survey respondents identified as having a current or past year

alcohol use disorder, and around 21% of people in the same survey with major depression also reported an alcohol use disorder over the previous 12-month period. Rates of comorbid SUDs are highest amongst people with bipolar disorders, with as many as 60% of people with bipolar I disorder reporting lifetime SUDs. In the United States, data indicate that people with cannabis dependence are twice as likely to have experienced clinically significant depression, and that current users are also twice as likely to have a concurrent mood disorder. Cannabis abuse has been associated with four times the risk of developing a subsequent depression. Generally, alcohol use disorders are most prevalent amongst people with mood disorders, followed by marijuana, amphetamines, and cocaine.

Service-based research suggests that levels of comorbidity and associated problems may be more pronounced in treatment-seeking populations. The ECA study reported that rates of AOD use disorders within mental health services were twice that of people with mental disorders not engaged in treatment, and rates of mental disorder were even higher among people engaged in AOD treatment. In general, mood disorders occur in 40–42% of people engaged in addiction treatment services. Depressive disorders can occur in as many as 60% of treatment seekers for cocaine dependence, as evidenced in a US-based survey of people recruited to research studies over a 16-month period, whereby 53% ($n = 184$) reported lifetime use of cocaine and of these, 89% met criteria for lifetime major depressive disorder.

ETIOLOGICAL MODELS OF ALCOHOL/OTHER DRUG USE AND MOOD PROBLEMS

Given the frequency with which mood and AOD use problems co-occur, population-based surveys have been used to more closely examine the various environmental, genetic, and neurobiological factors that may explain this particular type of comorbidity. As a result, several different models have been proposed to explain the co-occurrence of any mental health and AOD use problem and these can be applied to the relationship between mood and AOD use problems, with a view toward planning treatment approaches for these comorbid conditions. Figure 52.1 displays the following models:

1. Primary AOD use models propose that the problematic AOD use preceded the first problematic mood symptoms, and AOD is considered the major determinant of the symptoms experienced by the person. Often, mood symptoms would occur only within reasonably close proximity (1 month or less) to AOD intoxication, withdrawal or cessation. In this case, treatment would target the primary problem (AOD use), assuming that upon abstinence, the mood symptoms would also remit.
2. Primary mood models suggest that AOD use commenced after the development of mood problems, as a means to self-medicate symptoms, or to cope with distress related to symptom development. Mood symptoms would also be experienced in the absence of any AOD use, or over a sustained period (greater than 1 month) following AOD cessation. According to this model, the mood problem would be the target of treatment, under the assumption that relief of negative mood and associated distress would lead to a remission of AOD use.
3. Bidirectional models state that low mood acts as a trigger and maintaining factor for harmful AOD use, and vice versa.
4. Common-factor models suggest that one or more factors, such as genetic predisposition, antisocial personality disorder, social difficulties, stressful events, etc., contribute to an increased risk of both mood and AOD use problems.
5. Artifact models suggest that mood and AOD use problems are unrelated, and the high co-occurrence of these problems could simply represent sampling biases in population and/or clinical studies, or the overlap in diagnostic classification systems for major depression and AOD abuse, dependence, and withdrawal syndromes.

Little firm evidence exists to support the adoption of one of the above models over another, and each model has inherent problems when applied to individuals in real-world clinical settings. For example, while longitudinal studies of the comorbidity of marijuana use and depressive symptomatology support a primary marijuana use model, cross-sectional surveys indicate that a primary depression model better accounts for the dysphoria experienced with marijuana abuse in adulthood. Further, while the US-based ECA study revealed that rates of alcohol use disorders were 35% more likely in relatives of people with major depressive disorder, indicating a relationship between the two disorders, a second study found no such association.

We need large-scale follow-up studies of epidemiologically defined cohorts of adults to examine relationships between comorbid disorders. The recent 3-year follow-up of incident disorders in the NESARC cohort and their relationship to baseline disorders provides a good illustration of the potential value of this approach. In this study, 34 653 persons who were interviewed in the first wave of the NESARC study were followed up after 36 months and reassessed for

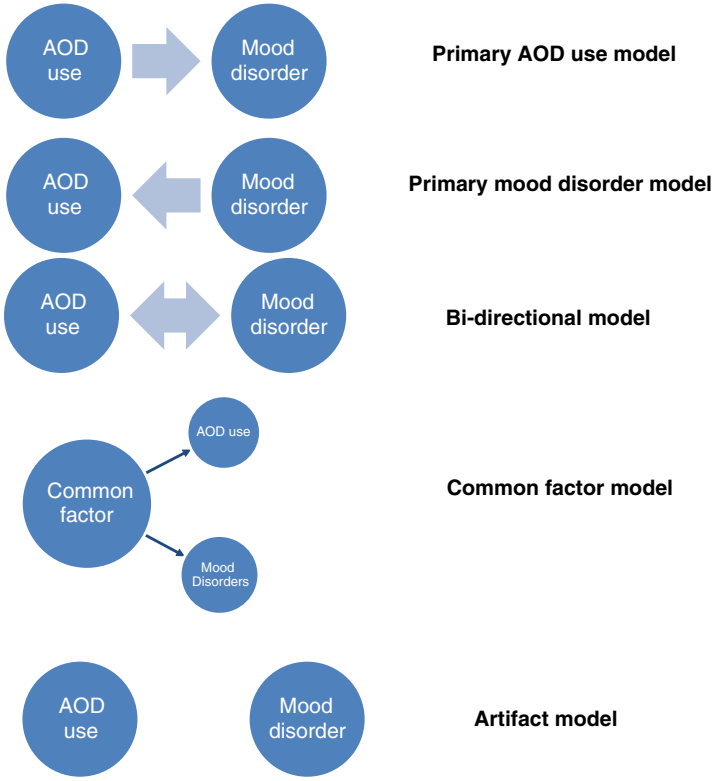


FIGURE 52.1 Proposed models of etiology of AOD and mood disorders.

symptoms of a wide range of common mental and SUDs. The sample size was large enough to examine relationships between demographic characteristics and mental disorders at the baseline assessment and incident disorders during the 3-year follow-up period. The disorders with the highest incidence rates (per 100 person years) were alcohol dependence (1.7), alcohol abuse (1.0), major depressive disorder (1.5), and generalized anxiety disorder (1.1). The demographic predictors of these disorders were unsurprising: alcohol and drug use disorders were more common among young males; and major depressive disorder was more common among young females.

The presence of some mental disorders at baseline did predict incident disorders during the 3-year follow-up (after controlling for demographic variables). Alcohol abuse and dependence were strong predictors of each other, as were drug abuse and drug dependence. By contrast, alcohol and drug use disorders at baseline did not predict an increased incidence of mood disorders; mood disorders at baseline, specifically bipolar and panic disorders, did predict an increased risk of drug and alcohol abuse at follow-up, as did borderline personality disorder.

The US NESARC did attempt to quantify the prevalence of self-medication among people with SUDs and mood disorders. Results indicated that one-quarter of the participants with mood disorders used alcohol/other drug to relieve their symptoms. The highest prevalence of self-medication was among people with bipolar I disorder, with 41% of participants indicating relief of symptoms as their primary reason for use of alcohol/other drugs. Further, men in this study were twice as likely as women to report self-medication of their mood disorder with alcohol/other drugs. In a 20-year study of mood and SUDs, researchers in the United States further revealed that the presence of manic symptoms in people with bipolar disorders was predictive of the development of subsequent alcohol use disorders, cannabis use disorders, and benzodiazepine use disorders. Bipolar II disorder was associated with later development of alcohol use disorders and benzodiazepine use disorders, with major depressive disorders associated with the later development of benzodiazepine use disorders. The self-medication model of substance use and mood disorder comorbidity assumes that a specific mood disorder would lead to abuse or dependence on a specific drug type that targeted the required symptom relief. The finding that, at least for bipolar disorders, multiple substances were associated with the presence of this disorder indicates that perhaps the self-medication model of comorbidity may not be an appropriate global etiological model of comorbid mood and SUDs.

Other researchers have suggested that the primary/secondary distinction between mood and AOD use disorders is immaterial once the two disorders have surfaced. This assertion is supported by studies indicating that the primary/secondary distinction, at least for alcohol use problems and depression, is not predictive of treatment outcomes. For example, in a 5-year follow-up of 97 Americans with concurrent affective disorder and alcohol use problems, levels of depression and alcohol use were equivalent for primary depressive and primary alcohol use conditions.

Clinical practice suggests that the relationship between AOD use and mood disorders may change over time. For example, depression may trigger alcohol use at some times and the reverse may occur at others. In a 10-year follow-up of people with alcohol-related problems, results indicated that there was a bimodal association between depression and transition to higher levels of drinking, perhaps indicating more than one developmental pathway. That is, respondents who progressed from low-level drinking to higher levels appeared to be drinking to relieve tension and distress (self-medicating their depression), while those who were already drinking at high levels at baseline seemed to have already developed chronic drinking patterns that contributed to their current depressive episode, rather than the reverse. This may be indicative of the clinical situation where initially, an individual may be drinking to relieve negative mood and other symptoms of depression (the depression is primary), however over time, the drinking may develop into a chronic problem in its own right, leading to depressive episodes. In this case, if treatment was based on the primary depressive condition, the independent alcohol use problem may not receive appropriate treatment.

Different etiological models may apply for men and women who experience comorbid mood and AOD use disorders. In a study of 2945 people with alcohol use problems, 15% were classified as having primary depression and a further 26% as having secondary (or alcohol-induced) depression. This study reported that there was a significantly higher proportion of men than women in the secondary depressed group (70% versus 30%), whereas the gender ratio was roughly equivalent for those in the primary depression group (48% men versus 52% women). Epidemiological studies additionally suggest that AOD use problems are more prevalent among men than women, with affective and anxiety problems more commonly reported among women. One implication of this pattern may be that it is more acceptable for males to report problems with drugs and alcohol than to admit to problems with their mood, emotions, or feelings. The reverse is likely to be true for females. This could lead to underreporting of affective states in men and of drug and alcohol use in women in survey-based research, further confounding the issues.

There are also practical problems in reliably establishing the temporal relationship between mood and AOD use problems, and using client self-report to establish the primary condition is often imprecise. This is especially true if both conditions have been present for many years, and if the primacy of each condition has changed over time. As such, a reliance on treating coexisting mood and AOD use problems according to the primary/secondary model may be inadequate. Further, the pattern of depressive symptoms in primary and secondary depression (coexisting with AOD use problems) is often so similar that it is difficult for clinicians to use the presence of certain key depressive symptoms as indicators of primary and secondary depressive conditions. Thus, clinicians may not be able to identify from the outset which clients are more likely to have a mood disorder that persists beyond abstinence from AOD use problems.

THE IMPACT OF COMORBID ALCOHOL/OTHER DRUG USE AND MOOD DISORDERS

Despite the difficulties in determining the etiological relationship between mood and SUDs, it is well known and accepted that people with this type of comorbidity experience significant and multiple adverse consequences of their comorbidity. For example, in one of the few AOD treatment studies that did not specifically exclude people with comorbid issues, a clinical profile of people presenting for treatment for marijuana use disorders was established. Results indicated that 83% experienced marijuana-related health problems, such as respiratory distress (e.g. asthma, coughing), problems with relationships, and psychological symptomatology (e.g. depression). Financial difficulties, criminal activity, and use of multiple drugs in addition to cannabis over their lifetime were high.

The severity, impact, and patterns of comorbidity of mood and SUDs are similar across international borders. Consistently, this type of comorbidity is associated with significantly more severe clinical symptoms, elevated suicidal behaviors, and poorer clinical course and outcome than for people without comorbidity. For example, compared with patients with bipolar disorder alone, those with bipolar disorder and substance use disorder comorbidity experience earlier onset of mood symptoms; higher rates of anxiety disorders, suicide attempts, accidents, hospitalizations, and rapid cycling; more depressive episodes; and lower treatment compliance. At a population level, the presence of comorbid depression and AOD use disorders has been associated

with high levels of disability, days out of role, and a significant contribution to the total burden of disease. In addition, epidemiological research suggests that there is a consistent association between suicidality, mood, and AOD use disorders. For example, in an analysis of the US Multiple Cause of Death Data, comorbid AOD use disorder with mood disorders (both depressive and bipolar) was associated with between 1.5 and 9.5 times the likelihood of death by suicide and other unnatural causes than for mood disorders alone. Data from this study also indicated that people with comorbid mood and SUDs died 12–34 years sooner than did people without this comorbidity.

Using data from NSMHWB, those respondents with comorbid alcohol use disorders and anxiety or depression experienced around twice as many days out of role in the previous 4-week period relative to their counterparts with alcohol use disorder alone. A higher proportion of the comorbid group than the alcohol use only group was younger, never married, unemployed, and separated or divorced. People in the comorbid group were four times more likely to have seen a specialist health professional for their problems, and slightly more likely to have seen any health professional in the previous 12 months. At entry to treatment, those participants with comorbid anxiety and depressive disorders reported more disability and higher levels of alcohol use than did those with alcohol use alone. Other research indicates that levels of depression are associated with higher rates of AOD use (such as tobacco) upon presentation to treatment services.

Service-based data indicate that, among people presenting for treatment for methamphetamine use problems, up to 70% concurrently experience at least moderate levels of depression, and that comorbid mood disorders are associated with significantly higher methamphetamine use, abuse and dependence, and significantly greater use of benzodiazepines, tobacco, and use of multiple substances.

A small body of research exists that attempts to examine the relationship between depression and AOD use and treatment utilization and outcome. The focus here has been almost exclusively on depression and comorbid alcohol use disorder. For example, people with comorbid depressive and alcohol use disorders are also suggested to require inpatient hospitalization more frequently than their counterparts without problematic alcohol use. In a large-scale US study of people with affective disorders, 3% also had a diagnosis of alcohol use disorder and reported increased rates of recurrent depressive episodes over the follow-up period, as measured by re-admission rates over 23 years.

The presence of mood problems at entry to treatment for alcohol dependence has also been associated with premature treatment dropout, more frequent relapses to alcohol use, and subsequent addiction treatment if left unaddressed. Antidepressant and other psychiatric medication can also be less effective when people are drinking alcohol at harmful levels. For example, in a study of people with major depressive disorder, levels of alcohol, tobacco, and caffeine use were assessed prior to commencement of an 8-week course of the antidepressant, fluoxetine. At the conclusion of treatment, levels of alcohol use at baseline were positively correlated with levels of depressive symptoms, suggesting a poorer response to the antidepressant.

Comorbid alcohol use and mood disorders additionally place people at increased risk of criminality. The same is likely true for drugs in addition to alcohol. For example, among a sample of 1016 methamphetamine users in the United States, levels of psychiatric disorder were high, with around one-third of the sample reporting depressive symptoms in the month prior to the study. Twenty-seven percent of the sample reported suicide attempts in their lifetime, and assault charges and other criminal activity were also high.

Despite these findings, other researchers suggest that the effect of a comorbid depressive disorder on AOD use characteristics is not clear-cut, as several researchers have failed to show a demonstrable effect of depression on AOD use and vice versa. For example, in a study of 54 undergraduate students in China, risk-taking tendencies were significantly reduced after induction of a negative mood, relative to positive and neutral moods. Although not directly tested in this study, one implication of this result may be that drug use, arguably a "risky" behavior, may be reduced in severity or frequency in the presence of depressive symptoms.

In a sample of 75 people with alcohol abuse/dependence, depressive symptoms present at the initial assessment were not related to quantity or frequency of drinking at baseline, nor to levels of alcohol consumption at the 3-month follow-up assessment. This study also suggested that depression may indeed enhance motivation to reduce drinking behavior, by increasing the alcohol-related costs/losses incurred by the individual. According to the stages of change model, behavior change will likely to occur when these personal costs for drinking outweigh any benefits the person perceives are associated with continuing to drink. Results of this study supported this hypothesis, with higher levels of depressive symptoms being significantly associated with the action stage of change for alcohol.

Levels of lifetime and current major depressive disorder were used to predict rates of relapse to problematic alcohol use following a brief intervention among 99 men with moderate to severe alcohol dependence. At the 6-month follow-up, 58% of the sample had relapsed to drinking. However, no differences existed between

relapsers and non-relapsers in terms of levels of depression at baseline, nor presence of lifetime major depressive disorder.

In summary, it is acknowledged that comorbid mood and SUDs present a set of unique difficulties and risks than would on either condition present on its own. However, little firm evidence exists to clearly suggest how these might apply to drugs in addition to alcohol, nor about the characteristics and needs of people within these populations who are seeking treatment. More research needs to examine these issues in more detail.

SMOKING AND MOOD DISORDERS

Results of the NSMHWB showed that 32% of current smokers had a mental disorder in the previous 12 months, which was twice the prevalence of 12-month mental disorders of those who had never smoked (16%). Compared to never smokers, current smokers also experienced four times the prevalence of 12-month SUDs (12% versus 3%), almost three times the prevalence of affective disorders (12% versus 5%) , and twice the prevalence of 12-month anxiety disorders (22% versus 11%). People with histories of depression or anxiety have been found to be heavier and more frequent smokers and have lower quit rates (although this is controversial) and higher dependence compared to people without such histories.

The relationship between smoking and mental disorders is thought to be complex, as mental disorders are a risk factor for smoking. Depressive mood increases the risk of smoking and vice versa. Commonalities in genetic and environmental factors probably underlie the comorbidity of depression and smoking.

SUMMARY

Mood and AOD use disorders commonly co-occur, and researchers and clinicians find it difficult to agree on how to best categorize and treat these concurrent conditions. This is a concern given that on a global scale, mental and AOD use disorders contribute 20% to the total burden of disease borne by society, and evidence from some studies indicates that comorbidity is associated with poorer functioning, higher risk of relapse, and increased symptom severity. When tobacco smoking is added to this scenario, the comorbidity between smoking, mood, and AOD use disorders becomes even more burdensome and associated with higher rates of mortality in addition to worsened morbidity.

SEE ALSO

International Data on the Prevalence and Correlates of Comorbid Substance Use and Psychiatric Disorders, Models of Relationships between Substance Use and Mental Disorders, Cannabis Use and the Development and Maintenance of Psychosis, Impact of Substance Use on the Course of Serious Mental Disorders, Substance Use in Response to Anxiety Disorders

List of Abbreviations

ATOD alcohol, tobacco, and other drug
AOD alcohol/other drug use
ECA study Epidemiology Catchment Area study, National Institute of Mental Health, United States, baseline conducted 1980–1985
NESARC National Epidemiological Survey on Alcohol and Related Conditions
NSMHWB National Survey of Mental Health and Wellbeing, Australian Institute of Health and Welfare, Australia, first conducted in 1997
SUDs substance use disorders

Further Reading

Abraham, H.D., Fava, M., 1999. Order of onset of substance abuse and depression in a sample of depressed outpatients. Comprehensive Psychiatry 40, 44–50.

Aihw, 1999. The Burden of Disease and Injury in Australia. The Australian Institute of Health and Welfare, Canberra.

Bovasso, G.B., 2001. Cannabis use as a risk factor for depressive symptoms. American Journal of Psychiatry 158, 2033–2037.

Brown, R.A., Inaba, R.K., Christian, G.J., Schuckit, M.A., Stewart, M.A., Irwin, M.R., 1995. Alcoholism and affective disorder: clinical course of depressive symptoms. American Journal of Psychiatry 152, 45–52.

Burns, L., Teesson, M., 2002. Alcohol use disorders comorbid with anxiety, depression and drug use disorders. Findings from the Australian National Survey of Mental Health and Well Being. Drug and Alcohol Dependence 68, 299–307.

Crum, R.M., Brown, C., Liang, K.-Y., Eaton, W.W., 2001. The association of depression and problem drinking: the results of the Baltimore ECA follow-up study. Addictive Behaviors 26, 76–773.

Grant, B.F., Goldstein, R.B., Chou, S.P., Huang, B., Stinson, F.S., Dawson, D.A., et al., 2008. Sociodemographic and psychopathologic predictors of first incidence of DSM-IV substance use, mood and anxiety disorders: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. Molecular Psychiatry Apr 22 [Epub ahead of print] doi: 10.1038/mp.2008.1041.

Hasin, D.S., Tsai, W.-Y., Endicott, J., Mueller, T.I., Coryell, W., Keller, M., 1996. Five-year course of major depression: effects of comorbid alcoholism. Journal of Affective Disorders 41, 63–70.

Kay-Lambkin, F.J., Baker, A., Lewin, T., 2004. The "comorbidity roundabout": a framework to guide assessment and intervention strategies and engineer change among people with co-morbid problems. Drug and Alcohol Review 23, 407–424.

Regier, D.A., Farmer, M.E., Rae, D.S., Locke, B.Z., Keith, S.J., Judd, L.L., et al., 1990. Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA)

Study. Journal of American Medical Association 264, 2511–2518 [see comments].
Patton, G.C., Coffey, C., Carlin, J.B., Degenhardt, L., Lynskey, M., Hall, W.D., 2002. Cannabis use and mental health in young pesssople: Cohort study. British Medical Journal 325, 1195–1198.
Schuckit, M.A., Tipp, J.E., Bergman, M., Reich, W., Hesselbrock, V.M., Smith, T.L., 1997. Comparison of induced and independent major depressive disorders in 2,945 alcoholics. American Journal of Psychiatry 154, 948–957.
Trosclair, A., Dube, S.R. Smoking among adults reporting lifetime depression, anxiety, anxiety with depression, and major depressive episode, United States, 2005-2006. Addictive Behaviors 35, 438–443.
Volkow, N.D., 2004. The reality of comorbidity: depression and drug abuse. Biological Psychiatry 56, 714–717.

Relevant Websites

http://www.dsm5.org/Pages/Default.aspx – American Psychiatric Association DSM-5 Development describes current diagnostic guidelines for the full range of mental disorders that will appear in the upcoming Diagnostic and Statistical Manual of the Mental Disorders (DSM-5) in May 2013.

http://www.mhfa.com.au – Mental Health First Aid provides guidelines on how to help a person developing a mental health or substance use problem.

http://www.ncpic.org.au – National Cannabis Prevention and Information Centre provides information for those experiencing problems with cannabis use, free resources, and training on related issues.

http://ndarc.med.unsw.edu.au/ – National Drug and Alcohol Research Centre website provides a range of research findings and related reports and papers, including links to information and online treatment programs for comorbid mood and substance use disorders.

http://www.Sane.org.au – Sane Australia is a national charity working for a better life for people affected by mental illness. It provides information and factsheets on mental health problems and substance use.

http://www.talktofrank.com/ – Talk to Frank website is an independent UK government-funded site. It includes information on a range of drugs, various personal stories, as well as a text-based question and answer service.

http://www.drug-rehabs.com/ – The Drug Rehab websites provides information and treatment resources for individuals suffering all addictions.

CHAPTER

53

Substance Use in Response to Anxiety Disorders

Jennifer A. Robinson, James M. Bolton

PsycHealth Centre, Winnipeg, MB, Canada

OUTLINE

Those who choose to use substances in response to an anxiety disorder typically do so in an attempt to ameliorate anxious symptoms. This behavior was first described in the scientific literature over 50 years ago, and was labeled the tension-reduction theory. Since then, a large body of research has accumulated to describe and define various aspects of this process currently known as self-medication. Edward J. Khantzian's formulation of the self-medication hypothesis (SMH) is based on two primary assumptions: (1) substance use is effective in reducing psychiatric symptoms and (2) the particular substance chosen for self-medication purposes corresponds to an individual's particular symptomatology. For instance, those with anxiety disorders are more likely to use alcohol to alleviate distressing symptoms than cocaine, given that alcohol is a central nervous system depressant and cocaine is a stimulant. Although evidence exists to support both of these claims, the former rests on a more solid empirical grounding than the latter.

A more recent formulation of the SMH makes three main assumptions: (1) psychiatric symptoms must precede substance use, (2) substance use is effective in relieving these symptoms, and (3) continued and excessive use of the substance is promoted by symptom relief. Support exists for each of these assumptions, but results

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00053-X

can vary depending on the particular anxiety disorder in question. Furthermore, examination of these processes has often been compromised by studies with methodological limitations. For example, many have been limited to cross-sectional samples, which prevent investigations of the central temporal claims of the SMH.

This chapter focuses specifically on the use of substances for the amelioration of anxiety symptomatology. Clinically and scientifically important questions to ask about this behavior include the following:

- What are the risks of self-medication?
- Who is most likely to engage in this behavior and when?
- Does substance use actually reduce anxiety? If so, how does it do this?
- How does this behavior affect treatment for anxiety?

Empirical investigations are ongoing, and in some cases these questions have only just begun to be examined. However, an attempt is made to highlight the major points in the body of knowledge that exists on the use of substances as an attempt to relieve anxiety (self-medication).

IMPORTANCE OF SELF-MEDICATION

Self-medication is an important and popular topic of scientific study, largely because this process is thought to (at least partially) account for the high comorbidity seen between substance use and anxiety disorders. Simply put, those who utilize substances in response to symptoms of anxiety are at higher risk for developing a comorbid substance use disorder, which can then lead to further mental health consequences. The prevalence and consequences of comorbidity are discussed as a basis for understanding the importance of clarifying the mechanism(s) behind this comorbidity.

Prevalence of Anxiety and Substance Use Disorder Comorbidity

Epidemiologic Studies

In the Netherlands Mental Health and Incidence Study, 13% of those with anxiety had comorbid substance use disorders, and 18.9% of those with substance use disorders had comorbid anxiety. Similar rates were found with the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). In this national survey of the US adult population, 15% of those with anxiety disorders had comorbid substance use and 17.7% of those with substance use had comorbid anxiety disorders. These rates were noticeably higher than the prevalence rates of the pure disorders found in the general population (which are approximately 9.4% for substance use disorders and 11.1% for anxiety disorders). Using the same sample, it was found that meeting criteria for a substance use disorder was associated with an almost twofold risk of also meeting criteria for an anxiety disorder.

Analysis of two more population samples – the National Comorbidity Survey and the Ontario Health Survey, from the United States and Canada, respectively – revealed that problem drinking was significantly associated with social phobia, simple phobia, agoraphobia, panic, and generalized anxiety disorder (GAD). A study that utilized four general population samples from different locations found that individuals with alcohol use disorders were more than twice as likely to receive a diagnosis of an anxiety disorder compared with those without an alcohol use disorder. These results indicate that substance use and anxiety disorders are substantial risk factors for each other and frequently co-occur in the general population.

With regard to specific anxiety disorders, among those with GAD in the general population, roughly half have a comorbid substance use disorder. Those with panic disorder and obsessive compulsive disorder are also significantly more likely to be diagnosed with an alcohol use disorder, and social phobia is consistently associated with high drug and alcohol use disorder comorbidity. One estimate reported that approximately 20% of those with social phobia also have an alcohol use disorder and 15% of those receiving alcoholism treatment met criteria for social anxiety disorder.

Clinical Studies

Comorbidity rates in clinical samples tend to be substantially higher than rates in general population samples. Among a sample of individuals receiving treatment for alcohol use disorders, 69% experienced anxiety-related problems. Among those being treated for anxiety disorders, an estimated 16–25% met the criteria for comorbid alcohol use disorders.

Individuals who present for treatment in a clinical setting may be experiencing a greater amount of psychological distress than those in the general population who are not currently seeking treatment. This increased severity may account for the higher comorbidity rates generally seen in treatment samples. Research in support of this confirms that comorbid anxiety and substance use are associated with increased rates of treatment seeking in a number of international locations.

Consequences of Anxiety and Substance Use Disorder Comorbidity

In both clinical and population samples, comorbidity has been associated with a number of serious negative health consequences that often exceed those of

individual or pure disorders. For instance, comorbid individuals tend to utilize mental health services at greater rates, have more severe symptomatology and disability, and have longer duration of illness compared with a noncomorbid cohort. These factors signify the importance of uncovering and studying mechanisms that can lead to this comorbidity in order to prevent its occurrence. Many researchers have attempted to investigate (using various methodologies and assumptions) whether self-medicating behavior is responsible for this comorbidity, as well as how and when it arises. These investigations are discussed subsequent to an exposition of the prevalence and consequences of self-medication in both epidemiologic and clinical samples.

Prevalence of Self-medication

Epidemiologic Studies

Prevalence rates for self-medication have been calculated using a variety of different samples, with results varying as a function of the particular anxiety disorder in question, sample characteristics, as well as the operationalization of terms such as anxiety and self-medication. A general population study calculated that self-reported self-medication was present in approximately 20% of respondents who met criteria for any anxiety disorder. This figure is roughly in accordance with other epidemiologic studies. Self-medication is typically assessed within these studies as a single self-reported item. In the National Epidemiological Survey on Alcohol and Related Conditions (NESARC), for example, after endorsing symptoms of panic disorder, the respondent is asked "did you ever drink alcohol to keep from having panic attacks?" Self-report items from other surveys are comparable with this.

Within the individual anxiety disorders, analysis of the National Comorbidity Survey revealed rates of self-medication ranging from 7.9% (social phobia, speaking subtype) to 35.6% (GAD). A different population study found the highest rates of self-medication with alcohol among those with social phobia (14.9%), with GAD following closely behind (14.1%). In a comparison of those with GAD and a substance use disorder with those with GAD and no substance use disorder, those with GAD and a substance use disorder were found to be 5 times more likely to self-medicate with alcohol and 14 times more likely to self-medicate with drugs.

Self-reported measures of self-medication may be prone to recall errors. More robust methods of assessing self-medication would involve assessment as self-medication was occurring, rather than asking about it after the fact. However, given its high prevalence in substance use and anxiety disorders, self-medication may play an important role in the maintenance of comorbidity. The use of drugs for self-medicating anxiety generally has a much lower prevalence than self-medication with alcohol, with rates ranging from about 1 to 5% among those with anxiety disorders.

Clinical Studies

Clinical samples are generally characterized by much higher rates of self-medication than community-based samples. In a sample of individuals with agoraphobia, 41% self-reported self-medication. Among those with comorbid alcohol abuse and agoraphobia, 91% self-medicated with alcohol. All of those with comorbid alcohol dependence and agoraphobia or social phobia said that alcohol was useful in coping with fears, and 33% intentionally used it for such purposes. In a different sample, 74% of socially phobic individuals reported using alcohol to reduce fears related to socialization. Given such high rates of self-medication with alcohol, particularly in response to social anxiety, it is not surprising that the empirical literature is focused on the comorbidity of these two disorders to the near exclusion of other substance and anxiety disorders. However, research is emerging to suggest that increased focus on other substances of abuse and other anxiety disorders (particularly GAD) may prove to be clinically and empirically relevant.

Consequences of Self-medication

In cross-sectional population-based studies, self-medication has been associated with extremely high odds of a number of different substance use disorders (such as alcohol, cocaine, stimulant, and marijuana dependence).

Besides potentially leading to comorbid substance use disorders, self-medication for anxiety may have a number of other negative health consequences. Cross-sectionally, this behavior is strongly associated with mood and personality disorders. Comparisons between self-medicating with alcohol and self-medicating with drugs have revealed that the risk of other comorbidities is higher among those self-medicating with drugs than with alcohol. Mental health-related quality of life is lower in those self-medicating with alcohol, and even lower among those self-medicating with drugs compared with those with anxiety disorders who do not self-medicate. Most concerning is that people who engage in self-medication are at increased risk of suicidal ideation and suicide attempts compared with those who do not self-medicate, indicating that this behavior may be a worrisome marker of safety concerns. These relationships hold even after adjustments are made for a range of potentially confounding variables, including comorbid

psychiatric diagnoses. Although it is known that comorbidity can have serious negative consequences (as described above), the results of these studies suggest that even in the absence of comorbidity, self-medication is an important clinical target for the prevention of important health consequences.

EVIDENCE FOR SELF-MEDICATION

Order of Onset

Researchers who support the premise that self-medication for anxiety disorders leads to a subsequent alcohol or substance use disorder must show that anxiety disorders precede substance use. This premise has proved somewhat difficult to investigate, largely due to the lack of prospective studies that have examined this question, combined with potential recall bias and the failure to examine different anxiety disorders separately. However, results seem to indicate that anxiety generally precedes alcohol or substance use disorders with, as we will see later, the possible exceptions of panic and generalized anxiety disorder.

A nationally representative dataset was examined to determine the retrospective age of onset of various disorders among comorbid individuals. Interesting differences emerged depending on the particular disorder examined. For instance, whereas both social and specific phobia tended to precede the onset of alcohol use disorders, panic and generalized anxiety disorder tended to follow them. The phobias were five times more likely to precede than to follow alcohol use disorders, with GAD four times more likely to follow alcohol use disorders. The findings regarding social phobia and panic accord fairly well with the established scientific literature on anxiety–substance use disorder comorbidity, whereas findings regarding GAD are at odds with some of the literature. One clinical study reported that 67% of comorbid cases of GAD and alcohol use disorders had an onset of GAD prior to the alcohol use disorder. This difference might be accounted for by the differences in sampling techniques. In other words, in the general population, GAD may be more likely to follow alcohol use disorder, whereas the reverse may be true in clinical samples. Further investigation on this issue is needed to clarify the order of onset for GAD and alcohol use disorder. In terms of panic disorder, there is evidence to suggest that the use of alcohol can condition and kindle panic attacks, supporting the finding that panic should more frequently follow than precede alcohol use disorders. Many studies consistently find that social phobia is a risk factor for the development of alcohol use disorders, which also fits with the above results.

In general, prospective studies show that baseline anxiety disorders present an increased risk for development of subsequent alcohol use disorders, and vice versa. A substantial amount of this longitudinal work has been unable to examine different anxiety disorders individually, and has focused primarily on alcohol use disorders as an outcome as opposed to other drug use disorders. A recent longitudinal study found that panic, social phobia, and GAD predicted the onset of substance use disorders, and substance use disorders predicted these same three anxiety disorders. All types of substance use disorders (alcohol abuse and dependence, and drug abuse and dependence) predicted the onset of GAD, whereas only alcohol dependence and drug abuse predicted the onset of panic disorder. However, this study did not examine mechanisms, such as self-medication, that might underlie this comorbidity.

Although it is possible that self-medication in response to anxiety disorders subsequently leads to the development of substance use disorders via increasing amounts of substance use to remedy anxiety, a slightly more complex, although inherently plausible model is also important to consider. This model is sometimes termed the feed-forward model, and suggests that self-medication of subthreshold anxiety symptoms can not only exacerbate substance use problems but can also increase anxiety symptomatology. Although immediately anxiolytic, over longer periods of time, substance use may become anxiogenic. One study of patients with phobias found that although the phobia typically preceded alcohol use disorders, it was most severe after alcohol use. Furthermore, withdrawal symptoms – especially withdrawal from alcohol – can mimic anxiety symptoms, leading to further self-medication. It is likely that these processes occur in many self-medicating individuals, and that continued self-medication for anxiety would promote the emergence of both disorders.

Anxiety reduction through extinction may also be prevented due to substance use. In other words, the use of alcohol prevents the experience of anxiety from occurring fully, which prevents learning that the situation is less frightening than imagined through exposure. This, combined with the negative reinforcement of immediate anxiety reduction through substance use, perpetuates both anxiety and substance use.

A major strength of this model is that it does away with questions about order of onset of disorders. Self-medication cannot be assumed simply because an anxiety disorder happens to have begun prior to a substance use disorder. Moreover, certain disorders are more likely to occur at different points along an individual's developmental trajectory simply due to the nature of the disorder. Overall, whether self-medication begins in response to subthreshold anxiety or to a full-blown anxiety disorder, it is likely that

among those with comorbidity, each disorder serves to reinforce the other.

Does Alcohol Reduce Anxiety?

Experimental studies aiming to examine whether alcohol actually decreases anxiety have been somewhat ambiguous up to this point. This can be contrasted with subjective reports among self-medicators, who typically report decreased anxiety as a result of alcohol use. Discrepancies within the experimental literature may be due to factors that were not accounted for, such as individual personality and sociodemographic characteristics. Therefore, it may be the case that alcohol use decreases subjective anxiety only in certain individuals. This notion has been supported by other studies examining factors potentially leading to alcohol and substance use in response to anxiety, and are discussed in more detail later. It has also been postulated that experimental studies that have not found an anxiety-reducing effect may not have allowed subjects to consume a sufficient amount of alcohol to generate the effect. Moreover, these studies have varied in the anxiety-provoking task given to participants. Differences seen in the effectiveness of alcohol to reduce anxiety may be the result of differences in baseline anxiety as a result of the particular task given.

In general, most experts agree that anxiety can be temporarily reduced by sufficient alcohol consumption, due to the pharmacological effects of ethanol, although this should not imply that alcohol use does not also increase anxiety (as delineated in the feed-forward model above). The effect of alcohol on anxiety reduction, as well as one's propensity to use it for these purposes, is likely to be the result of a number of factors including age, gender, expected effects, and other individual characteristics. In experimental studies that induce panic- and phobic-like symptoms, the use of alcohol has often been found to be anxiety reducing, although this is not the case for all studies.

Unequivocally, a large number of individuals with comorbid anxiety and substance use disorders report the use of substances as a means to cope with anxiety, and a large percentage of those with anxiety disorders report self-medication in both clinical and community samples. These subjective self-reports of anxiety reduction among self-medicating individuals is more consistent than reports in the experimental literature noted above. Anxiety reduction is the prominent affective consequence reported by self-medicators; the results from the experimental literature (which has examined those with and without propensities to self-medicate) are somewhat more ambiguous.

How Does Alcohol Affect Anxiety?

There are a number of potential mechanisms through which alcohol can have anxiolytic effects in anxious individuals. Alcohol affects the brain in ways that are similar to substances known to reduce anxiety, such as benzodiazepines and opiates. These substances operate by increasing the activity of γ-aminobutyric acid (GABA), a neurotransmitter that contributes to the inhibition of central nervous system activity. As many of the symptoms of anxiety are manifestations of increased autonomic arousal, the sudden muted activity in the nervous system may be perceived as a reduction of subjective anxiety.

Moreover, alcohol limits the production of norepinephrine, a substance that activates the central nervous system's fight or flight response, causing increases in blood pressure and heart rate. Another suggestion is that alcohol intoxication works to reduce anxiety by diverting attention away from anxiety-inducing stimuli, in what has been termed the attention allocation model. This prevents extinction from occurring and may lead to increased substance use as an avoidance strategy. Over 20 years ago, it was proposed that the expectation that alcohol will reduce anxiety leads individuals to self-medicate and to experience less anxiety, primarily due to a placebo effect. This notion has received increased attention in the research literature over the past decade or so and is discussed in greater detail below. Alcohol's anxiogenic effects often occur following a period of heavy drinking. While initially anxiety reducing, alcohol can often serve to increase anxiety in the longer term.

MEDIATORS AND MODERATORS

Alcohol Outcome Expectancies

Some findings indicate that alcohol itself does not decrease anxiety, but rather the mere belief that one has consumed it does. It is more likely that the belief that alcohol will reduce anxiety interacts with the anxiolytic depressant properties of ethanol. Alcohol outcome expectancies (AOEs) can be defined as an individual's evaluation of the anticipated effects of alcohol prior to its consumption. AOEs are a frequent topic of discussion in this body of scientific literature, particularly as they relate to social anxiety disorder. Because those who do not expect alcohol to reduce anxiety often report no anxiety reduction following drinking (at least in an experimental setting), AOEs may be important considerations when treating a client who is self-medicating for anxiety. Many reports indicate that the expectation that alcohol will reduce anxiety predicts increased consumption of alcohol, as well as the use of alcohol for the specific purpose of self-medicating symptoms.

Moreover, coping-related alcohol use is associated with greater problems related to alcohol use.

There is some evidence that among those who do not expect alcohol to reduce anxiety, those with higher levels of social phobia are less likely to consume alcohol compared with those with lower levels of social phobia who also do not expect alcohol to reduce anxiety. This may be the result of differences in avoidance behavior. For instance, those with higher social anxiety may be less likely to engage in social situations where alcohol is available. Conversely, those with lower social anxiety may be more likely to attend social gatherings, but to do so with fear and trepidation. Another possibility is that instead of anxiety reduction, these people may believe that alcohol will lead to disinhibition and greater risk for social embarrassment, which will therefore lead to less alcohol intake among those with more significant social phobia. Among those who did expect alcohol to reduce anxiety, no difference has been found between those with high or low social anxiety; both groups use alcohol for self-medication.

In general, the more a person believes that alcohol will reduce anxiety, the more likely they are to utilize it for such purposes. This belief may be maintained and perpetuated by popular notions that alcohol will "take the edge off" after a difficult day or help one "loosen up" at parties. Although this may be true, it can also lead to the development of substance use disorders and other psychiatric problems in vulnerable individuals. In considering the particular types of situations in which self-medication is most likely (which is examined in greater detail later), alcohol consumption is reduced in situations where performance may be impeded, such as before giving a public lecture. Despite its immediate anxiolytic effects, the awareness of increased anxiety subsequent to a perceived below par performance may prevent self-medicating in these specific situations.

Sociodemographic Variables

Gender

In the general population, men and women report roughly equal rates of self-medication with alcohol in response to anxiety, however, men are more likely to self-medicate with drugs. This is particularly important given findings that self-medicating with drugs may pose a greater risk of other comorbidities than self-medication with alcohol alone. The use of drugs to cope with anxiety can be considered a marker of clinical severity, particularly in individuals using illicit substances rather than misusing prescription medications.

In longitudinal studies, women with social phobia are at greater risk of developing subsequent alcohol use disorders than are men with social phobia, with family support moderating this relationship. Whether this gender difference exists in those with anxiety disorders other than social phobia remains to be seen. There are a number of possible explanations for this finding, including increased social acceptability of drinking in men, or greater physiological effects of alcohol in women. Further research is needed to clarify the underlying mechanisms at work.

Gender differences have also been found in terms of AOEs. Men with high AOEs experience greater anxiety reduction after being given a placebo alcoholic beverage than men with low AOEs. This difference was not seen in women. It may therefore be especially important to assess AOEs in male patients.

Age

Self-medication for anxiety is most prevalent in the 30–44 years age group, especially when considering self-medication with drugs. Those over the age of 65 years report very low rates of self-medicating behavior.

Marital Status

Married individuals make up the majority of the group self-medicating with alcohol (60%); the rates are roughly equal among those who are married, separated, or single when considering self-medication with drugs.

Education

In any type of self-medicating behavior (either with alcohol or drugs), the majority of self-medicators have attended college at minimum. Those who did not complete high school, or completed high school and did not pursue further education, reported less self-medication. This finding may be the result of sampling bias, because the majority of individuals in the sample were those who had some college education.

Income

Interesting differences have been reported in terms of self-medication and income. Higher income categories (US$60 000+) represent the majority of those who self-medicate with alcohol. Lower income categories (US$0–19 999) represent the majority of those who self-medicate with drugs.

To summarize, it appears that those self-medicating with alcohol, but not with drugs, may represent a higher sociodemographic bracket. These individuals tend to be married, educated, and have higher household income levels. This profile is somewhat contrary to that often seen among those with psychiatric disorders; such individuals tend to be representative of the lower income and educational categories, and they are often single. Whether or not this is the cause or consequence of mental illness is not a subject of debate in this chapter, we merely

wish to point out – particularly to clinicians working with self-medicating individuals – that a different sociodemographic profile may define this group.

Specific Situations

In general, people at risk for self-medication tend to drink less alcohol prior to a performance task, likely because they do not want to impede performance. However, these individuals are at risk for self-medication following the task, and in other types of anxiety-provoking situations. Self-medication has a greater probability of occurring in situations that the individual chooses to endure with fear rather than avoid, which may explain the dearth of research on self-medication for specific phobias, as well as the low prevalence of this behavior in populations of individuals with specific phobias. Simply put, those with a pure specific phobia may choose to – and be able to – avoid the feared stimulus rather than endure it, precluding any need or benefit from self-medication. In the few population studies that have looked at specific phobias, the results are consistent in that this disorder has the lowest prevalence of self-medication compared with the other anxiety disorders. Conversely, socially phobic individuals are less able – and may be less willing – to avoid interactions with people. They are more likely to choose to endure feared situations rather than avoid them and use alcohol to allow them to endure feared social situations. Those with GAD are also less able to cope with fears through avoidance strategies, given that they are generalized to a wide variety of situations, which may explain the high prevalence of self-medication in these individuals.

Among those with social phobia specifically, anxieties about being observed or being scrutinized (such as when performing or giving a speech) tend to predict the onset of alcohol use disorder over and above anxieties related to social interaction. However, other research has indicated that self-medication is more common in situations where an individual has to interact socially than in those where they must perform. From these somewhat contradictory findings, it can be concluded that although self-medicating for social interactions is more common, it is specifically drinking to cope with observation that puts one at risk for the development of an alcohol use disorder. More research is needed to clarify the specific conditions under which anxiety can lead to substance use disorders in order to improve specificity in identifying at-risk populations. Research on other anxiety disorders, such as those with high prevalence of self-medication (e.g. GAD), is also likely to have substantial clinical benefits.

In summary, those at higher risk of self-medication may be those who have high expectations that alcohol will reduce anxiety, are not about to engage in a performance task, and are male. Self-medication poses a risk of subsequent alcohol use disorders in those with social phobia, and particularly those with observation-based fears.

OTHER QUESTIONS ABOUT SELF-MEDICATION

Why Do Some Choose Self-medication and Others Do Not?

In one of the original formulations of the self-medication hypothesis, it was proposed that individuals who engage in this behavior have a reduced ability to tolerate negative affect. In other words, self-medicating made difficult affective states easier to bear for vulnerable individuals who were compelled to reduce symptoms immediately. Other individuals may experience a significant amount of anxiety, but for other reasons are more able or willing to tolerate it, thus reducing their risk of self-medication. It is also possible that those who choose to self-medicate may have more severe manifestations of anxiety disorders and comorbidities than those who do not. Simply meeting criteria for an anxiety disorder may not increase the risk of self-medication; other factors may need to be in place. These factors may include ones that have already been discussed such as male gender, decreased tolerance of anxiety, and specific feared situations. Others are likely to include elements such as poor coping skills, family history of substance use, low resilience, environmental factors, as well as the perceived stigma of seeking more conventional treatments (such as psychotherapy). At this early stage in the research, it is extremely difficult to determine who is at risk for self-medication and who is not among those with anxiety disorders. It is known that engaging in this behavior can be harmful and risky, but it is not entirely known why some choose this coping strategy over others. This is an extremely important area for future research to pursue in order to develop strategies for the prevention of comorbid anxiety and substance use disorders, as well as to prevent other detrimental consequences of self-medication.

How Does Self-medication Lead to Comorbidity?

Individuals who self-medicate are known to be at increased risk for the development of comorbid substance use disorders. This may occur as a result of anxiety and substance use exacerbating each other. Attempts to self-medicate with alcohol, for instance, may be negatively reinforced initially. However, as this behavior continues it is likely to cause more anxiety as a result of neurochemical effects on the brain as well

as failure to appropriately process, and gain exposure to, the root cause of the anxiety. This increased anxiety is then likely to lead to increasing attempts to self-medicate, and symptoms thus progress in an upward spiral. Recent studies indicate that among those who endorse self-medication, the amount of alcohol consumed moderates the relationship between social anxiety and incident alcohol use disorders.

GAPS IN THE LITERATURE

Although recent epidemiologic studies have aided in determining the prevalence and consequences of self-medication in the general population, and experimental studies have furthered our understanding of AOEs and efficacy of self-medication, there are many existing gaps in the scientific literature that need to be addressed.

First, the vast majority of the literature is focused on social phobia and alcohol use. This is somewhat justified given the high rates of self-medication with alcohol in social phobia, yet further research is needed to examine the impact of the use of drugs on those with social phobia. This includes the use of illicit drugs and the misuse of prescription medications. Self-medication with drugs, although not as common as self-medication with alcohol, is still a prevalent and potentially serious behavior in those with social phobia. This is especially important given the increased negative consequences of drug use compared with alcohol use.

Second, more attention should be focused on the use of substances to relieve anxiety in those with other anxiety disorders, specifically GAD (given the high prevalence of this behavior in GAD and its potentially serious consequences).

The amount of substances used to self-medicate is another important consideration for future investigations. This is especially important when considering findings showing that self-medication – in the absence of a substance use disorder and after adjusting for comorbidity – is associated with important negative health consequences (e.g. suicidality, decreased quality of life). In these findings, the amount of alcohol consumed may be driving the associations rather than whether it constitutes self-medication. Longitudinal research aimed at addressing the finer details of the temporal association between substance use and anxiety, as well as the perceived efficacy and outcome expectancies of self-medication will also likely prove to be a fruitful avenue for future investigations.

TREATMENT CONSIDERATIONS

There is recognition that many individuals may not present for treatment. Self-medication may be an ingrained behavior in some individuals that has developed over a long period of time. The person may believe it to be an effective coping mechanism and thus not feel the need for formalized treatment. Furthermore, some individuals may not even recognize their substance use behavior to be problematic, or their anxiety to be distressing. Thus, there will likely remain a large group of people who engage in self-medication who are beyond the reach of treatment efforts. This underscores the importance of public health campaigns directed at anxiety, drug and alcohol use, and their interplay. Enhancing understanding of these conditions at the societal level will hopefully expand the effectiveness of treatment efforts.

Conversely, given the substantial comorbidity and negative consequences associated with self-medication, increased treatment seeking may be observed in a subset of people. In reviewing the literature, it seems, at least preliminarily, that self-medication is associated with greater rates of treatment seeking compared with those with anxiety disorders who do not self-medicate, when examined in an epidemiologic sample. Those who self-medicate are more likely to have seen a professional (e.g. a counselor or psychiatrist) and to have been prescribed medications than they are to present in an emergency room or to have been hospitalized. Treatment providers therefore need to be aware of self-medication in order to effectively care for these individuals who do present for help.

General Principles

Perhaps the most important treatment consideration of self-medication is its recognition, as well as an understanding of its consequences. Specifically, treatment providers need to have a high index of suspicion of self-medication among individuals with anxiety symptoms and disorders, given its high prevalence in this population. Care providers should directly ask all people with anxiety whether they use alcohol or illicit drugs, or misuse prescription drugs, to reduce their anxiety. If they endorse self-medication, the treatment provider needs to recognize that this person is at greater risk of alcohol and drug use disorders (if they have not yet developed), psychiatric comorbidity, poor quality of life, and suicidal behavior, compared with their anxious counterparts who do not self-medicate.

Self-medication implies a preexisting anxiety disorder, and therefore clinical guidelines for the treatment of anxiety disorders should be followed. Likewise, if the person has developed a comorbid substance use disorder, established guidelines for alcohol and drug use disorders should be followed. A review of specific treatment approaches for anxiety disorders and substance use disorders is beyond the scope of this

chapter; this review focuses on treatment approaches specific to self-medication.

In recent years, there has been a greater emphasis on the importance of treating co-occurring disorders simultaneously, with a developing evidence base for this approach. Co-occurring treatment initiatives lead to better outcomes compared with approaches that separate the treatment of anxiety and substance use disorders. Self-medication patients should be referred to such programs if they exist in the region. If not, treatment providers should make every effort to align their treatment as closely as possible to co-occurring disorder treatment guidelines.

Medications: The Use of Benzodiazepines

Caution should be used in the prescription of benzodiazepines to individuals who are at risk for self-medication. Research indicates that these individuals tend to shift their usage of benzodiazepines from an as-prescribed to an as-needed basis, which can lead to addiction. In general, prescription of benzodiazepines should be confined to agents with the least amount of abuse potential wherever possible. If their benefit is deemed to outweigh the potential risks associated with their use, benzodiazepines with longer half-lives, dosed regularly rather than as-needed, may reduce the likelihood of misuse and dose escalation. It is important to realize that even if the use of benzodiazepines does not lead to a comorbid substance use disorder in at risk individuals, simply misusing these medications for self-medication purposes may be associated with some of the other harmful outcomes of self-medicating behavior. It is also important for clinical practitioners to be aware of the potential for harmful drug interactions when they are considering writing a prescription for an individual who has been self-medicating.

Psychotherapy

Psychoeducation and cognitive behavioral therapy are two psychotherapeutic modalities that are likely to be effective in the treatment of people who engage in self-medication. They both alert patients to the knowledge that experiencing anxiety, rather than numbing it or avoiding it through substance use, is the best way to overcome anxiety in the long run. Psychoeducation can help people identify their problematic coping strategies, as well as understand the negative consequences associated with self-medication.

AOEs may prove to be a fruitful target for cognitive behavioral therapy in individuals with anxiety disorders prone to self-medication, as well as in comorbid individuals. For instance, beliefs that alcohol will be anxiety reducing may be challenged (given that alcohol produces greater anxiety over the longer term). This is important, given findings that high AOEs predict increased use of alcohol for coping, and that many socially phobic individuals hold such beliefs. Cognitive restructuring regarding whether anxiety can be endured will likely also prove beneficial, since many self-medicating individuals are likely to believe that they must immediately reduce anxiety in order to avoid serious negative consequences. In general, thoughts that promote and maintain self-medication should be uncovered and challenged through cognitive behavioral techniques.

Behavioral experiments can be easily devised to help reduce anxiety as well as reliance on substances to deal with it. For example, during later stages of therapy, a patient might choose to attend a party and not consume alcohol in order to test out how much they actually need it to control their anxiety. This might work best in patients who do not concurrently meet criteria for a substance use disorder, because withdrawal symptoms may need to be dealt with primarily in these individuals. The effectiveness of cognitive behavioral therapy on self-medication requires further study, ideally in the form of randomized controlled trials.

SUMMARY

The use of substances in response to anxiety (self-medication) is a fairly common and risky behavior. It is associated with increased comorbidity and suicidality, as well as decreased quality of life. The majority of research in this area has focused on the use of alcohol in response to social anxiety disorder, but there is evidence to suggest that the use of other substances, and in response to other anxiety disorders, is an important topic for both clinicians and researchers to consider. AOEs, gender, and the particular anxiety disorder in question are important considerations for determining the risk of self-medication, but more work is needed to uncover other variables likely to be relevant. Systematic studies of treatment efficacy in self-medicating and comorbid individuals are missing from the research literature. Traditional cognitive behavior therapy and client-centered techniques are likely to be effective for self-medication but require formalized intervention studies. Given the significant consequences that self-medication can have on an individual, their families, and society at large (in terms of the economic burden of comorbidity), it is an extremely important topic for both clinicians and researchers.

SEE ALSO

Alcohol Use Disorders, Emotions and Addictive Processes, Interpersonal Factors and Addictive Disorders, Self-Medication, Stress and Addiction

List of Abbreviations

AOEs	Alcohol outcome expectancies
GAD	generalized anxiety disorder
NESARC	National Epidemiologic Survey on Alcohol and Related Conditions
SMH	self-medication hypothesis

Glossary

Benzodiazepines a category of psychiatric medications that increase the effect of the γ-aminobutyric acid (GABA) neurotransmitter, resulting in the reduction of anxious symptomatology.

Comorbidity the presence of two or more psychiatric disorders occurring in the same individual at a given time.

Cross-sectional measured at one time point.

Epidemiologic referring to large-scale population-based sampling methods used to determine health patterns in the general population.

Longitudinal measured over two or more time points.

Self-medication the use of alcohol or drugs as an attempt to reduce psychiatric symptomatology; drugs may be either illicit or prescription (if used in a manner other than as directed by a physician).

Self-medication hypothesis the proposal that self-medicating behavior is the primary causal explanation for co-occurring substance use and anxiety disorders.

Further Reading

Abrams, K., Kushner, M., Medina, K.L., Voight, A., 2001. The pharmacologic and expectancy effects of alcohol on social anxiety in individuals with social phobia. Drug and Alcohol Dependence 64, 219–231.

Bolton, J., Cox, B., Clara, I., Sareen, J., 2006. Use of alcohol and drugs to self-medicate anxiety disorders in a nationally representative sample. Journal of Nervous and Mental Disease 194, 818–825.

Carrigan, M.H., Randall, C.L., 2003. Self-medication in social phobia: a review of the alcohol literature. Addictive Behavior 28, 269–284.

Falk, D.E., Yi, H.Y., Hilton, M.E., 2008. Age of onset and temporal sequencing of lifetime DSM-IV alcohol use disorders relative to comorbid mood and anxiety disorders. Drug and Alcohol Dependence 94, 234–245.

Grant, B.F., Stintson, F.S., Dawson, D.A., Chou, P.S., Dufour, M.C., Compton, W., Pickering, R., Kaplan, K., 2004a. Prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders: results from the national epidemiological survey on alcohol and related conditions. Archives of General Psychiatry 61, 807–816.

Khantzian, E.J., 1997. The self-medication hypothesis of substance use disorders: a reconsideration and recent applications. Harvard Review of Psychiatry 4, 231–244.

Kushner, M.G., Abrams, K., Borchardt, C., 2000. The relationship between anxiety disorders and alcohol use disorders: a review of major perspectives and findings. Clinical Psychology Review 20, 149–171.

Kushner, M.G., Mackenzie, T.B., Fiszdon, J., et al., 1996. Relationship between alcohol problems and anxiety disorders. Archives of General Psychiatry 53, 264–270.

Menary, K.R., Kushner, M.G., Maurer, E., Thuras, P., 2011. The prevalence and clinical implications of self-medication among individuals with anxiety disorder. Journal of Anxiety Disorders 25, 335–339.

Merikangas, K.R., Mehta, R.L., Molnar, B.E., et al., 1998. Comorbidity of substance use disorders with mood and anxiety disorders: results of the International Consortium in Psychiatric Epidemiology. Addictive Behavior 23, 893–907.

Robinson, J., Sareen, J., Cox, B.J., Bolton, J., 2009a. Self-medication of anxiety disorders with alcohol and drugs: results from a nationally representative sample. Journal of Anxiety Disorders 23, 38–45.

Robinson, J.A., Sareen, J., Cox, B.J., Bolton, J., 2009b. Correlates of self-medication for anxiety disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Journal of Nervous and Mental Disease 197, 873–878.

Swendsen, J.D., Merikangas, K.R., Canino, G.J., et al., 1998. The comorbidity of alcoholism with anxiety and depressive disorders in four geographic communities. Comprehensive Psychiatry 39, 176–184.

Swendsen, J.D., Tennen, H., Carney, M.A., et al., 2000. Mood and alcohol consumption: an experience sampling test of the self medication hypothesis. Journal of Abnormal Psychology 109, 198–204.

Relevant Websites

American Psychological Association. http://www.apa.org/index.aspx.

American Psychiatric Association. http://www.psych.org/.

The Anxiety Disorders Association of America. http://www.adaa.org/.

Centre for Addiction and Mental Health. http://www.camh.net/.

CHAPTER

54

Cannabis Use and the Development and Maintenance of Psychosis

Wayne Hall[*], *Louisa Degenhardt*[$]

[*]The University of Queensland, Herston, QLD, Australia, [$]University of New South Wales, Sydney, NSW, Australia

OUTLINE

INTRODUCTION

Over the past several decades, epidemiological studies in developed countries have consistently found associations between regular cannabis use and psychotic symptoms and disorders in the general population. Clinical studies have also reported higher rates of heavy cannabis use among new and chronic cases of schizophrenia. In this article, we address the following important questions raised by these findings:

1. What do these associations and other evidence indicate about the role of cannabis use in the onset of serious psychiatric disorders like psychosis with symptoms of delusions, hallucinations, and cognitive and social impairment?
2. How does cannabis use affect the course and outcome of psychoses like schizophrenia?
3. Is there similar evidence that cannabis use may increase the risk of developing bipolar disorders?
4. What implications does this evidence have for responding to young people with psychoses who use cannabis; warning young people of the mental health risks of cannabis use; social policies toward recreational cannabis use by young people?

CANNABIS USE AND SCHIZOPHRENIA ONSET

The inference that cannabis use is a contributory cause of psychosis requires that there is an association between cannabis use and psychosis; that chance is an unlikely explanation of the association; that cannabis use preceded the psychosis; and that plausible alternative explanations of the association can be excluded.

There is now good evidence from prospective studies that cannabis use and psychosis are associated, that chance is an unlikely explanation of the association,

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00054-1

and that cannabis use often precedes the development of psychosis. The major challenge has been in excluding two other explanations of the relationship: (1) that it is due to other uncontrolled factors, such as other drug use (e.g. amphetamines and alcohol), or a genetic predisposition that increases the likelihood of developing schizophrenia and using cannabis and (2) that the direction of the causal relationship is from psychosis to cannabis; that is, that persons with psychoses use cannabis to self-medicate the symptoms of their disorder.

The first good evidence that cannabis use may precipitate schizophrenia came from a 15-year prospective study of cannabis use and schizophrenia in 50 465 Swedish conscripts. This study investigated the relationship between self-reported cannabis use at age 18 and a diagnosis of schizophrenia in the Swedish psychiatric case register during the next 15 years. Those who had tried cannabis by age 18 were 2.4 times more likely to receive a diagnosis of schizophrenia than those who had not. The risk of a diagnosis of schizophrenia increased with the number of times cannabis had been used by age 18 (compared with those who had not used cannabis) from 1.3 times among those who had used cannabis 1–10 times, 3 times for those who had used between 10 and 50 times, and 6 times for those who had used more than 50 times. These risks were substantially reduced after statistical adjustment for having a psychiatric diagnosis at age 18, and having parents who had divorced (as an indicator of parental psychiatric disorder) but the relationships remained statistically significant.

Zammit et al. reported a 27-year follow-up of the Swedish cohort study, covering most of the risk period for the onset of psychotic disorders. This study provided better statistical control of a larger number of potential confounding variables that included other drug use, IQ, and known risk factors for schizophrenia and social integration. It was also able to distinguish between cases that occurred in the first 5 years and later. Zammit et al. also found that cannabis use at age 18 predicted an increased risk of schizophrenia during the follow-up period in a dose–response way and these relationships persisted after statistically controlling for the effects of other drug use, other potential confounders, and psychiatric symptoms at baseline. The relationship was a little stronger during the first 5 years, probably reflecting the decline in cannabis use that occurs with age. They estimated that 13% of cases of schizophrenia could be averted if all cannabis use were prevented.

Van Os and colleagues reported a 3-year study of the relationship between self-reported cannabis use and psychosis in a community sample of 4848 people in the Netherlands. Subjects were assessed at baseline on cannabis and other drug use and psychotic symptoms were assessed using a computerized diagnostic interview and a diagnosis of psychosis validated by a telephone interview with a psychiatrist or a psychologist. A consensus clinical judgment was made on whether individuals with psychotic disorder needed psychiatric care.

Van Os et al. replicated and extended the findings of the Swedish cohort study. Cannabis use at baseline predicted the risk of psychotic symptoms during the follow-up period in individuals who had not reported psychiatric symptoms at baseline. There was a dose–response relationship between the frequency of cannabis use and the psychotic symptoms during the follow-up period, and these relationships persisted when they statistically controlled for the effects of other drug use. The relationship between cannabis use and psychotic symptoms was stronger in persons with more severe psychotic symptoms who were judged to be in need of psychiatric care. Those who reported any psychotic symptoms at baseline were more likely to develop schizophrenia if they used cannabis than were individuals who did not. Van Os et al. estimated that cannabis use accounted for 13% of the psychotic symptoms and 50% of the psychotic disorders that needed treatment.

The Swedish and Dutch findings were replicated in a 4-year follow-up of 2437 adolescents and young adults between 1995 and 1999 in Munich. The subjects were assessed at baseline on cannabis use and psychotic symptoms were assessed in early adulthood using the Computerised International Diagnostic Interview. They found a dose–response relationship between self-reported cannabis use at baseline and psychotic symptoms. Young people who reported psychotic symptoms at baseline were much more likely to report psychotic symptoms at follow-up if they used cannabis than were peers who did not have such a history.

Arsenault et al. reported a prospective study of the relationship between adolescent cannabis use and psychosis in a New Zealand birth cohort ($N = 759$), whose members had been assessed intensively on risk factors for psychotic symptoms and disorders since birth. Psychotic disorders were assessed according to DSM-IV diagnostic criteria, with corroboration from family members or friends. Psychotic symptoms were first assessed at age 11, *before* the onset of cannabis use. They also examined the specificity of the association between cannabis use and psychosis by conducting analyses of the effects of (i) other drug use on psychotic symptoms and disorders and (ii) cannabis use on depressive disorders.

Arsenault et al. found a relationship between cannabis use by age 15 and an increased risk of psychotic symptoms by age 26 that persisted after controlling for other drug use. The relationship was no longer statistically significant after adjustment for

psychotic symptoms at age 11, but this probably reflected low statistical power given the small number of psychotic disorders in the sample. The measurement of cannabis and other drug use was crude (viz, none, 1–2 times, and 3 or more times). There was no relationship between other drug use and psychotic disorders and between cannabis use and depression. Earlier onset of cannabis use was more strongly related to psychotic symptoms and there was a suggestion that there was a higher risk of psychosis among cannabis users who reported psychotic symptoms at age 11.

Fergusson, Horwood and Swain-Campbell reported on a longitudinal study of the relationship between cannabis dependence at age 18 and psychotic symptoms at age 21 in the Christchurch birth cohort in New Zealand. They assessed cannabis dependence using DSM-IV criteria and psychotic symptoms by 10 items from the Symptom Checklist 90. Fergusson et al. were able to adjust for a large number of potential confounding variables, including self-reported psychotic symptoms at the previous assessment, other drug use, and other psychiatric disorders. Cannabis dependence at age 18 predicted an increased risk of psychotic symptoms at age 21 years (RR of 2.3) that remained significant after adjustment for potential confounders (RR of 1.8). Fergusson and colleagues also examined the association between cannabis and psychotic symptoms until age 25 years, using structural equations modeling to account for observed and nonobserved confounding factors. The association between cannabis and psychosis was not explained by confounding factors, and the direction of the relationship was from cannabis use to the symptoms of psychosis rather than vice versa.

McGrath and colleagues conducted a similar analysis of an Australian birth cohort in Brisbane. They examined the association between the duration of self-reported cannabis use (collected at age 21) and the likelihood of reporting psychosis-related outcomes (as assessed by the Peters et al. Delusions Inventory) in a cohort of 3801 young adults. They found that the number of psychosis-related symptoms reported increased linearly with the duration of cannabis use and that the relationship persisted when controlling for potential confounders. They were also able to replicate these relationships within 228 sibling pairs in the sample, thereby controlling for the effects of unmeasured confounders (such as shared genetic risk and family background).

In all these longitudinal studies, there have been consistent associations between self-reported cannabis use in adolescence and psychotic symptoms in early adult life. There was, however, some uncertainty about the timing of the first cannabis use because participants were assessed yearly or less often about their cannabis use during the preceding year or years.

Verdoux and colleagues addressed this issue by studying the more fine-grained temporal relationships between cannabis use and psychotic symptoms using an experience sampling method. They asked 79 college students to report on their drug use and psychotic symptoms at randomly selected time points, several times each day over 7 consecutive days. Subjects made their ratings when prompted by randomly programmed signals to a portable electronic device. Participants included high-level cannabis users ($N = 41$) and some ($N = 16$) who had reported one or more psychotic symptoms in the past month in a personal interview. Users reported more unusual perceptions and less hostility in periods after cannabis use, and cannabis use was more strongly associated with unusual perceptions in vulnerable individuals.

Henquet and colleagues have recently replicated these results using a similar sampling method to study the relationship between cannabis use, mood, and psychotic symptoms over 6-consecutive study days in 42 patients with psychotic disorders and 38 normal controls who were regular cannabis users. They found that cannabis use predicted short-term increases in positive mood and longer-term increases in hallucinations in both groups. The patients were more sensitive to both these effects of cannabis than the controls.

IMPACT OF CANNABIS USE ON THE COURSE OF SCHIZOPHRENIA

Clinical reports suggest that patients with schizophrenia who continue to use cannabis have more psychotic symptoms, respond poorly to neuroleptic drugs, and have a worse clinical course than those patients who do not. Prospective studies have reported similar findings. Jablensky and colleagues, for example, reported that patients with schizophrenia who used cannabis and cocaine had more psychotic symptoms and hospitalizations in a 2-year follow-up of 1202 first episode schizophrenic patients in a 10-country WHO Collaborative study. Martinez-Arevalo et al. reported that continued use of cannabis during a 1-year follow-up of 62 DSM-diagnosed schizophrenic patients predicted a higher rate of relapse and poorer compliance with antipsychotic drug treatment.

Linszen and colleagues reported a prospective study of 93 psychotic patients whose symptoms were assessed monthly over a year. Twenty-four of their patients were cannabis users (11 were less-than-daily users and 13 were daily users). Cannabis users relapsed to psychotic symptoms sooner, and had more frequent relapses in the year of follow-up, than the patients who had not used cannabis. Daily users relapsed earlier, and more often, than the less-than-daily users who, in turn, relapsed sooner, and more often, than the patients who did not

use cannabis. These relationships persisted after statistically controlling for premorbid adjustment and alcohol and other drug use during the follow-up.

These results were replicated in a larger sample of patients with schizophrenia followed by Degenhardt and colleagues monthly over a 10-month period. There was a small association between cannabis use in 1 month and psychotic symptoms in the following month that persisted after controlling for baseline symptoms and other drug use. Foti et al. have recently confirmed these results in a 10-year follow-up of 162 incident cases who were interviewed during their first admission, and then 6 months, 2, 4, and 10 years later. Cannabis use in the previous year predicted an increased likelihood of psychotic symptoms that persisted after adjustment for confounders. There was also some evidence that the experience of psychotic symptoms increased the likelihood of cannabis use in the following year.

Some uncertainty remains about the role of confounding factors, such as premorbid differences between patients who do and do not use cannabis in personality, family history, and other characteristics that predict a poorer outcome. The other challenge has been in separating the contributions that cannabis, alcohol, and other drug use may make to any exacerbation of schizophrenic symptoms. A 2008 meta-analysis of 13 longitudinal studies by Zammit and colleagues concluded that most studies were underpowered statistically and had not adequately controlled for baseline differences in the severity of disorder or for the effects of alcohol and other drug use.

INTERVENTION STUDIES

Intervention studies that reduce cannabis use among patients with schizophrenia who use cannabis could reveal whether symptoms of their disorders improved. The major difficulty with this strategy is that we have not been very successful in treating cannabis use disorders in persons with schizophrenia. There are very few controlled outcome studies of cannabis abuse treatment in schizophrenia and these studies have showed modest treatment effects and have rarely treated a large enough number of patients to detect any positive impacts of abstinence from cannabis on the course of schizophrenia (*see* chapter Treatment for Co-occurring Substance Abuse and Mental Health Disorders for more details).

BIOLOGICAL PLAUSIBILITY

The major psychoactive ingredient of cannabis, delta-9-tetrahydrocannabinol (THC), acts upon a specific cannabinoid receptor (CB_1) in the brain. D'Souza and colleagues have also shown in a double-blind provocation study that intravenous THC provokes positive and negative psychotic symptoms in a dose-dependent way in healthy volunteers and persons with schizophrenia. There is increasing evidence that the cannabinoid system may be involved in schizophrenia and related disorders. For example, CB_1 receptor knockout mice show behavior consistent with some of the symptoms of schizophrenia, such as reduced goal-directed activity and memory for temporal representations. Elevated levels of anandamide, an endogenous cannabinoid with similar effects to that of THC, have also been found in the cerebrospinal fluid of persons with schizophrenia.

AN OVERALL EVALUATION

The evidence reviewed points strongly in the direction of cannabis use playing a causal role in the precipitation and exacerbation of schizophrenia and other psychoses. Two other explanations need to be excluded before drawing these conclusions.

The first is that young people with psychoses use cannabis to self-medicate the symptoms of their disorders. The evidence for his hypothesis is weak. The reasons that most persons with schizophrenia give for using alcohol, cannabis, and other illicit drugs are similar to those given by peers who do not have schizophrenia, namely, to relieve boredom, provide stimulation, feel good, and socialize with peers. They also use the same drugs as their peers: tobacco, alcohol, and cannabis. Some persons with schizophrenia report using cannabis to relieve negative symptoms and depression, but prospective epidemiological studies have generally not found any relationship between early psychotic symptoms and an increased risk of later cannabis use, as required by the self-medication hypothesis. Moreover, neither Verdoux and colleagues nor Henquet and colleagues found a temporal relationship between reporting unusual experiences and the use of cannabis, as would be the case if self-medication was involved.

The other major alternative to a causal explanation of the association between cannabis use and psychosis is that it is due to uncontrolled residual confounding. The most plausible confounders are (1) the use of other drugs that can produce psychotic symptoms, such as, psychostimulants and alcohol and (2) a genetic vulnerability to develop a psychosis that also increases the risk of using cannabis or the susceptibility to its effects. The strongest epidemiological studies have addressed these forms of confounding by statistically adjusting for other drug use, personal characteristics that predict psychosis risk, and a personal history of psychotic symptoms. The number of confounding variables that have been assessed has varied between studies, as have the specific

variables that have been statistically controlled for. One recent study used fixed effects regression to control for *unmeasured* confounders.

In most of these studies, the relationship has persisted after adjustment for potential confounders. We accordingly agree with Moore and colleagues in their review of the literature that it is unlikely that the association is due to confounding. Those who continue to argue for uncontrolled confounding need to identify *plausible* confounding variables that have not been adequately controlled in studies to date so that they can be controlled in future studies. We also give weight to the evidence on the biological plausibility of a causal relationship between cannabis use and psychosis. For these reasons, we think it is more likely than not that cannabis use precipitates schizophrenia in persons who are vulnerable because of a personal or family history of schizophrenia. This hypothesis is consistent with the stress-diathesis model of schizophrenia that schizophrenia arises from the effects of stress acting on individuals with a preexisting vulnerability (arising from family and personal history) to develop the disorder.

This explanation is also consistent with the following facts: the relative risk of developing schizophrenia among regular cannabis users is modest (RR of around 2–3); the incidence of treated schizophrenia does not appear to have substantially increased during the period when cannabis use has steeply increased among young adults in Australia, North America, and the United Kingdom, and the onset of these disorders is, on average, at an earlier age among those who have used cannabis.

It is also noteworthy that heavy alcohol and amphetamine use have been accepted as causing psychoses on the basis of weaker evidence than that for cannabis. For example, the evidence that heavy alcohol use causes psychosis largely consists of case series of *delirium tremens* in severely alcohol-dependent people undergoing alcohol withdrawal. There is one experimental study, which showed that *delirium tremens* could be induced by abruptly stopping alcohol after several weeks of sustained heavy drinking in a hospital ward.

The evidence that heavy amphetamine use can induce a psychosis was initially based on Connell's 1958 series of 200 case studies of heavy amphetamine users who developed paranoid psychoses after sustained heavy amphetamine use and whose disorders remitted after abstinence. These case series were supported by small-scale experimental reproductions of psychoses in amphetamine users and normal volunteers. More recently, epidemiological studies have found associations between the frequency of amphetamine injection and the severity of psychotic symptoms in illicit amphetamine users. A causal explanation is supported by animal evidence that amphetamine and cocaine affect dopaminergic neurotransmission.

CANNABIS USE AND BIPOLAR DISORDER

Like schizophrenia, bipolar disorders also often have begun in late adolescence and early adulthood during the peak periods for initiation of cannabis use. Persons with bipolar disorders also have high rates of comorbid cannabis use disorders and there is some evidence that heavy cannabis use may also precipitate symptoms of these disorders. Henquet and colleagues reported a 3-year longitudinal study of the relationship between self-reported cannabis use and symptoms of mania in the NEMESIS sample of 4848 people in the Netherlands. Cannabis use at baseline predicted an increased risk of manic symptoms during the follow-up period in individuals who had not reported symptoms at baseline; there was a dose–response relationship between the frequency of cannabis use at baseline and the risk of manic symptoms during the follow-up; and these relationships persisted when they statistically controlled for the effects of personal characteristics and other drug use.

There is similar evidence for an adverse impact of cannabis use on the course of bipolar disorders. A North American study by Strakowski of 144 bipolar patients, 69 of whom had cannabis use disorders, found poorer outcomes in cannabis users but this difference disappeared after adjustment for confounders. The effects of cannabis use on outcome were also much less pronounced than the effects of alcohol use disorders in the same cohort. Different results were reported in a European study of the outcomes of bipolar disorder after a 12-month follow-up in 3459 patients. In this study, cannabis users showed poorer treatment compliance, higher levels of overall symptoms of disorder, and of mania and psychotic symptoms; and all of these relationships persisted after controlling for demographic and clinical confounders.

Overall, the evidence suggests that there are similar relationships between cannabis use and bipolar disorder as between cannabis use and schizophreniform psychoses, namely, early and regular use increases manic symptoms, and patients with bipolar disorders who are regular cannabis users are more likely to experience symptoms of the disorder. The evidence is not as strong as that for schizophrenia, but is reasonable enough to warrant similar policy responses, a topic to which we turn now.

WHAT ARE THE POLICY IMPLICATIONS OF THE EVIDENCE?

A Public Health Case for Prudence

How strong does the evidence for a causal relationship between cannabis and psychosis need to be before

we are justified in taking action? If the standard of proof required for action was "beyond reasonable doubt," then we would find it difficult to take *any* public health policy decisions. If, however, we are prepared to act on the "balance of probabilities" (more likely than not), then policy action can be justified on prudential grounds.

In deciding on a policy in the face of uncertainty, we need to consider the likely costs and benefits of different courses of actions. Prudential reasoning arguably supports the efforts to discourage young people from using cannabis. The public health gain from discouraging use, if the relationship is causal (perhaps a 10% reduction in schizophrenia incidence), would arguably offset the foregone pleasure among those young people who either did not use cannabis, or delayed using cannabis until young adulthood. In principle, a reduction in cannabis use among incident cases of psychosis would also provide evidence for the effectiveness of this policy but it has been difficult to decide if the increased use of cannabis by young Australians has increased the incidence of schizophrenia. Recent epidemiological modeling in the United Kingdom by Hickman and colleagues also suggests that a very large number of young people would need to be prevented from becoming heavy cannabis users (2018–4530) in order to prevent a case of schizophrenia. Four to five times as many young people would need to be deterred from even light cannabis use in order to prevent one case of schizophrenia.

The case for attempting to discourage cannabis use by young people is strengthened by the evidence that regular cannabis use in adolescence predicts the development of cannabis dependence, poor educational outcomes, an increased risk of using other illicit drugs, an increased risk of depression, and a lower quality of life and poorer social relations in early adulthood. There are similar debates on the causal interpretation of these associations, but the fact that cannabis use is associated with so many indicators of poor psychosocial outcomes makes it unlikely that confounding explains them all.

Responding to Cannabis Use among People with Psychosis

The implications of the evidence are probably least controversial for mental health services that treat young people with psychoses, among whom there are high rates of regular cannabis use. Given the evidence that people with psychoses who are regular cannabis users have more positive symptoms, more frequent relapses, and require more hospitalization, it is prudent to encourage young people with psychotic disorders who use cannabis to stop or reduce its use. The major challenges are finding ways of persuading persons with schizophrenia to stop doing something that they enjoy and helping those who want to stop using cannabis but find it difficult to do so.

Recent evaluations of psychological interventions for cannabis dependence in persons *without* psychoses have found only 20–40% abstinence rates at the end of the treatment and there are substantial rates of relapse thereafter. The treatment substantially reduces cannabis use and problems among those who do not succeed in quitting. Many persons with schizophrenia have characteristics that predict a poor treatment outcome, namely, they lack social support, are cognitively impaired, unemployed, and may not adhere to treatment. A Cochrane review of 25 RCTs found that there was no compelling evidence to support one psychosocial treatment over another in treating people with serious mental illnesses who used cannabis. This finding was confirmed in a recent review by Baker and colleagues who identified CBT and motivational interviewing as the most promising approaches to explore.

INFORMING YOUNG PEOPLE ABOUT THE MENTAL HEALTH RISKS OF CANNABIS USE

A major challenge is finding effective ways of persuading young people that cannabis use is a contributory cause of psychosis and that regular cannabis use can increase poor psychosocial outcomes. This task is complicated by the polarized views on the risks of cannabis use expressed in the public policy debates about whether cannabis use by adults should continue to be a criminal offense. Publicly expressed differences of opinion about the evidence make it easier for young people to discount the evidence of harm.

We need to be realistic about the likely impacts of health education on drug use. Well-conducted school-based drug education can produce statistically significant reductions in cannabis use, but these effects are modest. The primary effect is on knowledge rather than behavior change and this is more likely to occur among less frequent users rather than the heavier users who are at greater risk of psychotic symptoms and other adverse effects. The nature and delivery of the advice will need to differ for different groups facing different levels of risk. The best way to deliver the advice will depend upon good research on the preexisting views of young people and the most effective ways to change them.

Education on the risks of cannabis use should be part of the general health education on drug use and mental health. It should explain the mental health risks of regular intoxication with alcohol and cannabis; and define the known high-risk groups such as those with

a family history of psychosis and those who have had bad experiences with cannabis and alcohol. It also needs to equip young people to identify peers who are adversely affected by cannabis use and to encourage them to cease using or seek help.

A major challenge will be framing the magnitude of the psychosis risks from cannabis use. If cannabis use increases the incidence of psychosis among those who use it regularly, then the risk for regular cannabis users increases from around 7 in 1000 to 14 in 1000, arguably still a low incidence rate. If this risk is multiplicative with family history, then in persons with an affected first-degree relative, the risk could increase from 1 in 10 among those who do not use cannabis to 1 in 5 or 1 in 3 among those who use cannabis. The consequences for those individuals who develop the disorder are serious. The temptation for parents and health educators is to play up the risk, arguing that everyone is at risk because it is difficult to predict which young people are most vulnerable. We think this a doubtful strategy that may undermine the credibility of the message by being seen to exaggerate the risk.

POLICIES TOWARD RECREATIONAL CANNABIS USE

The evidence on cannabis and psychosis is clearly relevant in societal decision making about cannabis policy because psychoses are serious disorders that adversely affect the lives of the young people affected by them. Nonetheless, this health effect cannot and should not be the sole basis for social policy. It does not follow that cannabis use should be prohibited because it harms some users. If it did, we would be morally obliged to prohibit alcohol and tobacco use, not to mention motor cars and motorbikes. Those who advocate for cannabis prohibition need to provide additional arguments that criminal penalties are the best way to discourage cannabis use and reduce the harms that it causes. As a society, we also need to consider the social costs of using the criminal law to deter people from using cannabis; that is, informed policy requires evidence on the harms caused by cannabis use, and the social consequences of policies that aim to discourage its use, including criminal penalties for possession and use (*see* more detailed discussion in Drug Decriminalization and Legalization on legalisation).

SUMMARY

There is consistent epidemiological evidence of a causal relationship between cannabis use and psychosis, namely, the two disorders co-occur in the population and clinical populations; cannabis use often precedes psychotic symptoms in prospective studies, and the association does not appear to be explained by plausible confounding variables. A causal relationship is also biologically plausible because the cannabinoids interact with dopaminergic neurotransmission, the cannabinoid system may be disturbed in psychosis, high doses of THC provoke psychotic symptoms in people without psychosis, and persons with psychoses who continue to use cannabis have more severe symptoms and poorer outcomes than those who do not.

We should accordingly discourage cannabis use among young people who have developed mental health problems. This could be done by screening all patients with psychotic symptoms, and advising those who use cannabis to stop, or at the very least, to reduce their use. More research is needed on (1) how best to persuade them to stop and (2) better ways of assisting those who would like to stop, but find it difficult to do so.

There is an ethical imperative to inform young people of the probable mental health risks of cannabis use, especially early and frequent use of cannabis. The challenge will be finding credible and persuasive ways of doing so.

We should avoid making the common assumption that if the relationship between cannabis use and psychosis is causal, then we should continue to criminalise cannabis use. Accepting a causal relationship weakens the simplest case for liberalization – the absence of any harmful effects of cannabis. Given the seriousness of the psychotic disorders for the life chances of the young people who are affected by them, this evidence increases the case for caution in liberalizing cannabis laws in ways that may increase young people's access to cannabis, decrease their age of first use, or increase their frequency of cannabis use. But the effect of the law on young people is not the only outcome we should consider in framing cannabis policies. A considered decision about a policy toward cannabis requires an analysis of the social harms caused by current policy, as much as the harms caused by cannabis use.

SEE ALSO

Marijuana Use and Abuse

Glossary

Bipolar disorder a serious mood disorder defined by one or more episodes of abnormally elevated mood (mania) and depression. Either type of episode can be accompanied by psychotic symptoms such as delusions and hallucinations.

Cannabinoids they comprise drugs derived from the cannabis plant, such as THC or CBD, and the synthetic substances that act on cannabinoid receptors in the brain.

Cannabis any product of the *cannabis sativa* plant, namely, marijuana (the flowering tops of the plant) and hash, the compressed resin.

Cannabis dependence marked distress resulting from impaired control over cannabis use as indicated by difficulty in ceasing use and continued, usually daily use, despite harms caused by such use.

Delusions fixed and false or fanciful beliefs that often involve being persecuted by others or having special talents and attributes. These beliefs are firmly sustained despite not being believed by anybody else and in the face of strong disconfirming evidence. They are not accepted by other members of the person's culture.

Hallucinations perceptions that occur in the absence of appropriate stimuli in a conscious and awake state. They are vivid and perceived as veridical perceptions by the person experiencing them. They can occur in any sensory modality, but auditory and visual hallucinations are most common.

Provocation study an experimental study in which participants (e.g. persons with or without schizophrenia) are exposed to a putatively causal substance (e.g. cannabis) or placebo to see if the substance provokes a specific response (e.g. psychotic symptoms).

Psychotic symptoms symptoms of serious mental illness (such as schizophrenia and bipolar disorder) that include hallucinations and delusions.

Schizophrenia a serious mental illness that typically begins in early adulthood. It is characterized by impaired thinking, loss of contact with reality, and emotional unresponsiveness. Among its most common symptoms are auditory hallucinations, delusions, disorganized speech and thought, blunted emotional expression, and significant social and occupational impairment.

Schizophreniform psychosis a psychotic disorder that shares some of the symptoms of schizophrenia such as delusions and hallucinations.

Tetrahydrocannabinol (THC) the chemical substance in cannabis plants that is primarily responsible for the psychoactive effects sought by cannabis users, namely, euphoria, relaxation, increased sociability, and so on.

Further Reading

Baker, A., Hides, L., Lubman, D., 2010. Treatment of cannabis use among people with psychotic or depressive disorders: a systematic review. Journal of Clinical Psychiatry 71, 247–254.

Degenhardt, L., Hall, W., Lynskey, M.T., 2003. Testing hypotheses about the relationship between cannabis use and psychosis. Drug and Alcohol Dependence 71, 37–48.

Degenhardt, L., Tennant, C., Gilmour, S., et al., 2007. The temporal dynamics of relationships between cannabis, psychosis and depression among young adults with psychotic disorders: findings from a 10-month prospective study. Psychological Medicine 37, 927–934.

D'Souza, D., Perry, E., MacDougall, L., et al., 2004. The psychotomimetic effects of intravenous delta-9-tetrahydrocannabinol in healthy individuals: implications for psychosis. Neuropsychopharmacology 29, 1558–1572.

ElSohly, M., Ross, S., Mehmedic, Z., et al., 2000. Potency trends of delta(9)-THC and other cannabinoids in confiscated marijuana from 1980–1997. Journal of Forensic Sciences 45, 24–30.

Green, B., Young, R., Kavanagh, D., 2005. Cannabis use and misuse prevalence among people with psychosis. British Journal of Psychiatry 187, 306–313.

Hall, W., Degenhardt, L., 2009. The adverse health effects of nonmedical cannabis use. Lancet 374, 1383–1391.

Henquet, C., van Os, J., Kuepper, R., et al., 2010. Psychosis reactivity to cannabis use in daily life: an experience sampling study. British Journal of Psychiatry 196, 447–453.

Linszen, D., Dingemans, P., Lenior, M., 1994. Cannabis abuse and the course of recent-onset schizophrenic disorders. Archives of General Psychiatry 51, 273–279.

Moore, T., Zammit, S., Lingford-Hughes, A., et al., 2007. Cannabis use and risk of psychotic or affective mental health outcomes: a systematic review. Lancet 370, 319–328.

Room, R., Fischer, B., Hall, W., Lenton, S., Reuter, P., 2010. Cannabis Policy: Moving Beyond Stalemate. Oxford University Press, Oxford, UK.

van Os, J., Bak, M., Hanssen, M., et al., 2002. Cannabis use and psychosis: a longitudinal population-based study. American Journal of Epidemiology 156, 319–327.

Verdoux, H., Gindre, C., Sorbara, F., Tournier, M., Swendsen, J., 2002. Cannabis use and the expression of psychosis vulnerability in daily life. European Psychiatry 17, 180S.

Zammit, S., Allebeck, P., Andréasson, S., Lundberg, I., Lewis, G., 2002. Self reported cannabis use as a risk factor for schizophrenia in Swedish conscripts of 1969: historical cohort study. British Medical Journal 325, 1199–1201.

Zammit, S., Moore, T., Lingford-Hughes, A., et al., 2008. Effects of cannabis use on outcomes of psychotic disorders: systematic review. British Journal of Psychiatry 193, 357–363.

Relevant Websites

http://www3.interscience.wiley.com – Addiction: A leading journal in the addictions field that publishes articles on cannabis.

http://www.emcdda.europa.eu – European Monitoring Centre on Drugs and Drug Addiction (EMCCDA): A leading European agency that publishes research on patterns of drug use and related harms.

http://www.nida.nih.gov – National Institute on Drug Abuse: A leading US agency for research on cannabis and other drugs.

http://www.unodc.org – United Nations Organization on Drugs and Crime: The leading UN agency on illicit drugs that publishes an annual World Drug Report.

http://www.who.int – World Health Organization (WHO): The leading UN health agency on substance abuse.

CHAPTER

55

Impact of Substance Use on the Course of Serious Mental Disorders

*Mark Deady**, *Maree Teesson**, *Kathleen T. Brady*[$]

[*]National Drug and Alcohol Research Centre, University of New South Wales, Sydney, NSW, Australia

[$]Medical University of South Carolina, Charleston, SC, USA

OUTLINE

CAUSES OF COMORBIDITY

While comorbidity between substance use disorders (SUDs) and severe mental disorders is common, the interaction between disorders is generally seen as one of mutual influence. There are a number of possible explanations for the frequent co-occurrence of mental disorders and SUDs. The explanations have differing consequences for the impact of SUDs on severe mental disorders and we cover them below. The causes include:

- The presence of a mental disorder may lead to an SUD, or vice versa (known as the direct causal hypothesis).
- There may be an indirect causal relationship (i.e. the role of an intermediary causal agent).
- There may be factors that are common to both the substance use and mental disorder, increasing the likelihood that they will co-occur.

Direct Causal Hypothesis

The Substance Use Disorder May Be a Consequence of the Mental Disorder

In some cases of substance use in individuals with mental disorders, the substance use occurs in an attempt to relieve mental health symptoms. If this happens repeatedly, over time an SUD is likely to develop. This is often described as the "self-medication hypothesis," in that substances are used in an attempt to medicate mental health symptoms. In these

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00055-3

circumstances, mental health conditions may become more apparent after the substance use has ceased. Certain mental health conditions may also impair a person's ability to make sound judgments regarding his or her substance use. For example, individuals with some personality characteristics or cognitive impairments relating to a mental disorder may have difficulty identifying social cues about appropriate use. This may lead the person to use in greater quantities or with greater frequency, increasing the likelihood of developing an SUD.

The Mental Disorder May Be a Consequence of Substance Use

Alternatively, substance intoxication and withdrawal can induce a variety of mental health symptoms and disorders, such as depression, anxiety, and psychosis. For example, alcohol use and withdrawal can induce symptoms of depression or anxiety; manic symptoms can be induced by intoxication with stimulants, steroids or hallucinogens; and psychotic symptoms can be induced by withdrawal from alcohol, or intoxication with amphetamines, cocaine, cannabis, Lysergic acid diethylamide (LSD) or Phencyclidine (PCP). In the majority of cases, these effects subside and eventually disappear with abstinence. For some, however, psychiatric symptoms may continue even after they have stopped drinking or using drugs.

Indirect Causal Relationship

An indirect causal relationship is said to exist if one condition has an effect upon an intermediary factor that, in turn, increases the likelihood of developing the second condition. For example, research has shown that the presence of early onset of substance use reduces the likelihood of completing high school, entering tertiary education, and completing tertiary education. This poor level of education may lead to later life difficulties (e.g. unemployment) that may predispose an individual to the development of other problems, such as depression. Similarly, the reverse is possible, whereby a mood disorder may lead to difficulties in completing study and work commitments, which may in turn lead to difficulties finding employment, increasing the risk of substance misuse.

Common Factors

The co-occurrence of two conditions may also come about due to the presence of shared biological, psychological, social, or environmental risk factors. That is, the factors that increase the risk of one condition may also increase the risk of another. For example, both substance and mental health conditions have been associated with lower socioeconomic status, cognitive impairment, the presence of conduct disorder in childhood, and antisocial personality disorder. It is also possible that a genetic vulnerability to one disorder may increase the risk of developing another disorder.

THE NEUROBIOLOGY OF SUBSTANCE USE IN SEVERE MENTAL DISORDERS

A growing body of evidence from basic science and translational studies implicates common neurobiologic pathways and abnormalities involved in addiction and a number of mental disorders. Using a neurobiologic framework, several hypotheses can be postulated to explain comorbidity: (1) addiction and other mental disorders are different symptomatic expressions of similar pre-existing neurobiologic abnormalities; (2) repeated drug administration, through neuroadaptation, leads to biologic changes that have common elements with the abnormalities mediating certain mental disorders.

One of the bridging constructs between mental disorders and SUDs involves the role of stress in SUDs and other psychiatric disorders. Corticotrophin releasing factor (CRF), one of the key hormones involved in the stress response, has been implicated in the pathophysiology of anxiety, affective and addictive disorders. Stress stimuli that activate CRF circuits are also known to potentiate mesolimbic dopaminergic reward pathways in laboratory animals. Similarly, human laboratory studies have shown that emotional stress and negative affect states increase drug craving in drug-dependent individuals. Evidence of an association between severity of depressive symptoms in patients with major depression and the subjective reinforcing effect of an acute dose of dextroamphetamine suggests dysregulation of reward systems with increasing level of distress in major depression. In this case, the addition of substance use to major depression has been shown to lead to a poorer course for the depression. The mechanism for this poorer course is dysregulation of reward systems. Animal models of early life stress and chronic stress result in long-term changes in stress responses which can alter the sensitivity of the dopamine system to stress and increase susceptibility to self-administration of substances of abuse. This may provide the neurobiologic underpinnings of the well-established relationship between early life adversity and SUDs in adolescents and adults and for the well-documented poorer course for severe mental disorders in those with SUDs

DOES CAUSALITY MATTER?

In the past, there has been a focus on establishing the relative order of onset of mental disorders and SUDs to identify which is the primary disorder. As, indicated earlier in this chapter, disorders may occur in any order, or they may develop at the same time. The evidence regarding the typical order of onset of disorders is not consistent. Establishing the order of onset of conditions can be useful in understanding the relationship between conditions. It is important to note however, that once both a mental disorder and substance misuse have been established, it is most likely that the relationship between them is one of mutual influence rather than there being a clear causal pathway, and both conditions may serve to maintain or exacerbate each other. For example, a person may engage in substance use to reduce symptoms of depression, however, research suggests repeated use may lead to increased depression. It is also possible that the relationship between disorders may change over time. For example, depression may trigger alcohol use on some occasions, while it may be the result of alcohol use on others.

WHAT ARE THE COMMON COMORBID DISORDERS AND WHAT IS THEIR IMPACT?

For many individuals with severe mental disorders and a comorbid disorder, a poorer disorder course is often the consequence. It is also the case that some substance use is more common among some mental disorders. Before examining course for each disorder, we will briefly outline the common comorbidities so as to set the research on course in perspective.

The most informative methodology for examining course of disorder is a longitudinal cohort study. These are, however, rare in individuals with severe mental disorders. The main source of information is long-term follow-up from clinical populations, usually following treatment. Research on the course of disorders, therefore, is very heavily reliant on treatment outcome rather than natural course.

SCHIZOPHRENIA

Prevalence

It has been approximated that 40–50% of those with schizophrenia will, at some point in their lives, also be diagnosed with an SUD; however, rates vary tremendously in published studies from as low as 10% to as high as 70%. The Epidemiologic Catchment Area (ECA) study of individuals from both community and institutional settings found 47% of all persons with a lifetime diagnosis of schizophrenia or schizophreniform disorder met criteria for some form of substance abuse or dependence (33.7% for alcohol disorder and 27.5% for another drug abuse disorder). Persons with schizophrenia were 4.6 times more likely to have a substance abuse diagnosis compared to the rest of the population (three times more likely for alcohol disorders, and six times more likely for other drug disorders). The moderating effects of demographic (e.g. gender, age) and population (e.g. clinical or nonclinical, in- or out-patient treatment) characteristics, along with other factors such as geographic location and heterogeneity in diagnosis/assessment techniques, are often cited as the reasons for such variability. Individuals with schizophrenia are, nevertheless, more likely to use substances than members of the wider community. As with many other forms of co-occurring conditions, the highest rates of comorbidity tend to be among those in treatment, the homeless, and the incarcerated.

Individuals with schizophrenia in particular have extremely high rates of smoking, ranging from 58 to 90%, with most studies suggesting it is almost universal. In addition, a recent review of 42 studies from 20 countries found that heavy smoking and high nicotine dependence were more frequent in smokers with schizophrenia than in smokers in the general population.

Clinical data suggest that cigarette smoking in individuals with schizophrenia may represent an attempt to self-medicate some symptoms of the illness. Studies suggest that individuals with schizophrenia have a primary deficit in brain nicotinergic systems that leads to abnormal sensory gating, and smoking cigarettes may be an attempt to overcome this deficit. Consistent with this theory, several studies indicate particular neuropsychological performance deficits in schizophrenia appear to be improved by nicotine administration and argue this may be one of the reasons for the high rates of nicotine smoking within this population.

With the exception of nicotine dependence, there is little evidence to suggest that persons with schizophrenia choose specific drugs to reduce specific symptoms and negative states. Other than nicotine, alcohol tends to be the drug most frequently associated with abuse or dependence in this population, followed by cannabis (and in some regions, cocaine). In general, the rates of types of substance use vary and tend to follow general population and geographic trends.

There has been much investigation and discussion surrounding the link between cannabis and psychosis. It is generally accepted that although there is little evidence to suggest that cannabis use per se causes psychosis, it appears that the drug exacerbates the

illness and may precipitate an earlier onset of illness in vulnerable individuals. In a recent meta-analysis of data from 83 studies, researchers found that cannabis users experienced psychosis almost 3 years earlier than nonusers, while for those with broadly defined substance use this onset was 2 years prior to that of nonusers. Cannabis and cocaine use have also been linked to psychotic relapses, while the role of alcohol is less clear. It has been suggested that individuals who use cannabis to ameliorate distressing symptoms of schizophrenia (rather than those for whom cannabis might be a predisposing factor to the disorder) report less negative symptoms than non-abusers. There may, however, be alternate explanations which, may better explain this finding.

Course

In assessing the impact of substance misuse on the course of mental disorders, there is a lack of clarity in the literature for a number of reasons, including inconsistencies in information concerning the primacy of disorders, restriction of samples to clinical populations, and differing treatments and rates of service use.

Overall, the impact of substance use on the course of schizophrenia appears to be detrimental across a range of outcomes. Generally persons with schizophrenia and comorbid substance use disorders have more psychotic symptoms than those without substance use. This link with psychosis may be mediated by treatment noncompliance among those with SUDs, as a number of reports indicate that persons with schizophrenia who abuse substances are more likely to be medication noncompliant. This medication noncompliance may be one explanation for the increase in psychotic symptoms.

There have been many studies highlighting the high rates of relapse and readmission to hospital and medication/treatment noncompliance among this comorbid population. Indeed, reports indicate that a decrease or cessation of substance use significantly increases the probability of remission from psychosis, while persistent use substantially reduces this likelihood. Although it is fair to assume these individuals with both schizophrenia and SUDs have a more complex clinical profile, recent evidence on treatment outcomes for people with co-occurring disorders indicates greater optimism is warranted, at least during the first 1–2 years. In the past, attempts to treat the disorders sequentially were largely unsuccessful and arguably to blame for these poorer outcomes. Currently, an integrated approach to treatment is recommended and integrated treatment programs have been rapidly emerging. Evidence suggests that if consistently applied, individuals in integrated treatment programs achieve positive recovery and more stable outcomes across substance use and schizophrenia as well as improved social and emotional outcomes and quality of life (*see* Treatment for Co-occurring Substance Abuse and Mental Health Disorders).

Recent studies into long-term recovery outcomes for this group are equally encouraging. At 3-year follow-up, one study found that participants with comorbid substance use and schizophrenia improved steadily in terms of reduced alcohol and drug use, schizophrenia symptoms, hospitalization and homelessness, as well as increased employment outcomes, contact with non-substance abusers, and overall life satisfaction. At 10-year follow-up, each of these areas showed steady and often significant improvement with the most dramatic functional improvements in the area of competitive employment. Participants also reported greater satisfaction with their lives overall. While most participants had not achieved the clinical cutoffs taken as indicators of recovery in several areas at the 3-year follow-up assessment, by 10 years a majority had achieved these levels on at least four of six recovery outcomes.

In terms of pharmacotherapeutic interventions in those with comorbid schizophrenia and SUDs, some studies suggest first-generation (typical) antipsychotic medications are less helpful than in non-comorbid individuals. There are even several reports suggesting that substance use may increase in such individuals when these antipsychotics are used. There is evidence, however, that the newer, second-generation (or atypical) antipsychotics (e.g. Clozapine) may be highly efficacious in individuals with schizophrenia and co-occurring SUDs and may have a positive impact on substance use in some individuals. Nevertheless, these benefits must be weighed against the heightened metabolic side effects of these medications.

Individuals with schizophrenia and substance use also tend to have increased mortality and higher rates of serious infection and medical problems, such as HIV and hepatitis C when compared to those without use. Because of the high rates of smoking among individuals with schizophrenia, this population is also at higher risk for many tobacco-related diseases compared to the general population. With these increased complications surrounding treatment response, relapse, and physical illness it is not surprising that hospitalization and crisis care expenses are also higher in this group when compared to individuals with schizophrenia without drug and alcohol problems.

Comorbid substance use has also been associated with a range of psychosocial problems. As with individuals with SUDs in general, those with schizophrenia and SUDs have been found to be at higher risk for violent behavior, suicide, and legal issues. One study found that individuals with schizophrenia and polysubstance use were more than 12 times more likely to be violent during a 3-month period when using and more than

four times more likely to be violent in the subsequent 3 months when compared to nonusers. Interestingly, cannabis or alcohol use alone did not predict violence. Consequently, these comorbid individuals are more likely to be incarcerated, while conversely, they are also more likely to be victims of violence themselves.

Issues surrounding stability of accommodation and homelessness have been shown to be more pronounced in this comorbid group compared to individuals with schizophrenia without SUDs. This comorbid group has also been found to benefit less from community outreach programs for the homeless. High rates of homelessness and unstable accommodation are arguably linked to the finding that substance use is associated with low levels of satisfaction with family relationships among persons with severe mental illness.

MAJOR DEPRESSIVE DISORDER

Prevalence

Like all prevalence data, a range of methodological, demographic, and population factors lead to variation in findings regarding the rates of SUDs in individuals with major depressive disorder. In the general population however, rates of current SUDs in this group range from 8.5 to 21.4%, with a lifetime prevalence of comorbid SUDs ranging from 27 to 40%. The ECA study found 27.2% of all persons with a lifetime diagnosis of major depression met criteria for some form of SUD (16.5% for alcohol disorder and 18% for another drug abuse disorder). Persons with depression were 1.9 times more likely to have a substance abuse diagnosis compared to the rest of the population (1.3 times more likely for alcohol disorders, and 3.8 times more likely for other drug disorders). Similarly, the National Comorbidity Study (NCS) indicated that those with major depressive disorder were twice as likely to be dependent on alcohol, twice as likely to be dependent on drugs and 1.6 times more likely to abuse drugs than those without major depression. The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) found that, among those in the general population with major depressive disorder, 19.2% had an SUD in the prior 12 months, 14% had a 12-month alcohol use disorder, and nearly 5% had a 12-month drug use disorder, while among those with lifetime major depressive disorder, this rate was 40 and 17% for alcohol and drug use disorders, respectively.

In clinical populations, the prevalence of concurrent SUDs in those with major depressive disorder tends to be higher than the general population. This is consistent across mental disorders as the presence of multiple disorders increases the likelihood of seeking treatment. In clinical populations, the rate of 12-month SUDs ranges from 8.6 to 25% and the lifetime prevalence ranged from 30 to 42.8% of treatment-seeking patients with major depressive disorder.

The types of SUDs most commonly found in those with major depressive disorder generally mirror those of the general population, with particularly high rates of alcohol use disorders. A recent meta-analysis of 35 studies revealed a median prevalence of current alcohol problems in depression of 16% (ranges 5–67%) and lifetime rates of 30% (ranges 10–60%). In contrast, the general population figures of these disorders were 7% for current and 16–24% for lifetime alcohol problems.

Course

The course of major depressive disorder in those with an SUD tends to be more complex than in those without this comorbidity. Recent epidemiological data has shown that among populations with lifetime drug or alcohol dependence and major depressive disorder, past year substance dependence remission was associated with an approximate three-fold reduced risk of depression. Compared with those without the comorbid condition, individuals with major depressive disorder and concurrent SUDs are more likely to be younger, male and either divorced or never married. They also tend to attribute greater functional impairment to their illness. Similarly, it has been reported that compared with those with major depressive disorder alone, those with comorbid SUDs have greater depressive symptomatology, more frequent concurrent anxiety disorders, poorer functioning (particularly social functioning), and an earlier age of onset of depression. Despite this finding, there have been no randomized controlled trials regarding comorbidity interventions targeting this younger population. A longitudinal study of disorder course in adolescents found that although an increase in substance use is not necessarily accompanied by an increase in depressive symptoms, decreases in substance use severity are associated with decreases in depressive state. Similarly, stable substance use rates, either at a low or a high level, tend to be associated with low or high levels of depression, respectively.

Much of the literature on course of major depressive disorder is found in the treatment literature. Longitudinal follow-ups of depressed patients indicate poor long-term outcomes for severe depressive disorders. These chronic outcomes are compounded by co-occurring substance use. Treatment outcome studies have consistently found that individuals with co-occurring major depressive disorder and SUDs have poorer treatment outcomes than those with one disorder alone, however, the nature of this relationship is contentious.

Among opiate-dependent populations, current mood disorders are associated with poorer treatment response and higher rates of relapse. Analogous findings have been observed among alcohol- and cocaine-dependent patients. The presence of an alcohol use disorder has also been shown to reduce the likelihood of depression remission, while the immediate treatment of the SUDs can reduce depressive symptoms. Meanwhile, a recent meta-analysis reported mixed findings regarding the effects of alcohol problems on depression course in terms of risk of relapse and likelihood of recovery, the authors conclude that current, but not lifetime, alcohol problems increase health care utilization. More generally, individuals with a range of comorbid SUDs and major depressive disorder have been shown to use health services more frequently and at greater cost compared to persons with one disorder. All these findings indicate both the negative impact comorbid substance use has on the course of depression and the positive effect associated with the treatment of these disorders.

There is growing evidence that antidepressants improve depression outcomes in persons with SUDs. Some studies have also found that antidepressant treatment can also improve drug use outcomes in comorbid individuals, but these effects are strongest in those individuals who report improvement in the symptoms of depression, suggesting that the improvement in SUDs outcomes may be influenced by the improvement in depression. It is generally observed that pharmacological intervention is not a stand-alone treatment, and concurrent therapy directly targeting the addiction is also indicated.

A recent meta-analysis has found only limited evidence for the effectiveness of cognitive behavioral therapy (CBT) either alone or in combination with antidepressant medication for the treatment of co-occurring depression and substance use, over and above that of other psychotherapies. There is, however, consistent evidence of improvements in both depression and substance use outcomes, regardless of the type of treatment. The authors suggest more specific forms of CBT be developed for this comorbid population. In light of this finding a number of integrated cognitive behavioral therapies have been developed with promising results both on substance use and depression outcomes.

Perhaps unsurprisingly, considering the high rates of suicide in the disorders respectively, individuals with major depressive disorder and a comorbid SUD have been found to have higher rates of suicidal ideation and attempts than in individuals without this comorbidity. Furthermore, evidence suggests, at least with alcohol use disorders, that this may be true of not only concurrent comorbidity, but also those with any history of alcoholism. Conversely, individuals with major depressive disorder and current suicidality have been shown to have higher rates of current SUDs than those without.

Evidence suggests that interpersonal and social functioning is often impaired in depressed patients with comorbid SUDs. Reports indicate that compared with individuals with major depressive disorder, only those with co-occurring alcohol problems have impaired relationships with their spouses, as well as a higher rate of divorce and living alone. Less evidence exists concerning the impact of other drugs on interpersonal and social impairment in major depressive disorder, however, limited data suggests similar patterns of poorer functioning.

BIPOLAR DISORDER

Prevalence

SUDs are common in individuals with bipolar disorder. Three epidemiologic studies conducted in the United States over the last 30 years have all concluded that individuals with bipolar disorder are at greatly enhanced risk for the development of SUDs as compared to the general population. In the ECA study, 56% of individuals with bipolar disorder had a lifetime SUD. The NCS found that individuals with bipolar disorder were approximately seven times more likely to have an SUD when compared to the general population. In both the ECA and NCS, bipolar disorder was the Axis I disorder associated with the highest risk for a co-occurring SUD. More recently, the NESARC also found that bipolar disorder was highly and significantly related to SUDs. Studies in clinical samples also support the high co-occurrence of bipolar disorder and SUDs.

There is little evidence to suggest that individuals with bipolar disorder choose specific drugs depending on their mood states. Alcohol is the drug most frequently associated with abuse or dependence in this population. In general, the rates of types of substances used by individuals with bipolar disorder vary and tend to follow general population and geographic trends.

Course

A number of studies suggest that the presence of an SUD has an impact on the age of onset and course for individuals with bipolar disorder. While several studies have found that the onset of bipolar disorder at an early age is associated with co-occurring SUDs, others have noted a later onset of bipolar disorder in alcohol-dependent bipolar patients. It has been hypothesized that there are subtypes of individuals with bipolar-SUDs comorbidity. The early onset group has more severe

bipolar illness that includes the development of subsequent SUDs, whereas the patients with SUDs that precede the onset of their mood disorder may have a less severe illness that required the presence of substance abuse to initiate the bipolar disorder. Several studies have found that mixed manic states and rapid cycling bipolar disorder are more common in bipolar individuals who have SUDs than in those who do not. This is important because both mixed episodes and rapid cycling disorder are associated with treatment resistance and a slower time to recover from an index episode. Along these lines, several studies have found that individuals with co-occurring bipolar disorder and SUDs have increased health service utilization including more emergency room visits and hospitalizations, more suicide attempts, and greater medication noncompliance.

Longitudinal studies have indicated that even a remitted SUD may be associated with poorer acute treatment response, a longer time to remission of their acute mood episode, and a greater percentage of time with subthreshold, but clinically significant, depression and manic symptoms over follow-up compared to those without this comorbidity. This poorer course of illness could not be fully explained by subsequent substance abuse during follow-up. However, a recent study found past or current SUD did not predict time to recovery from a depressive episode relative to no substance use comorbidity. Rather, those with current or past SUD were more likely to experience a switch from depression directly to a manic, hypomanic, or mixed state. These contrasting findings may be the result of differing forms of therapy or severity of substance use.

Much of the evidence on course of disorder in those with bipolar disorder is from long-term follow-up of treatment samples. Lithium is often considered the standard medication treatment for bipolar disorder, however, co-occurring SUDs may be a predictor of poor response to lithium. There is some data suggesting that anticonvulsant mood-stabilizing agents may be a better choice in individuals with SUDs. In particular, one recent trial investigating lithium plus valproate as compared to lithium plus placebo in subjects with alcohol dependence and bipolar disorder found better alcohol-related outcomes in those who received valproate. There have also been some promising pilot studies investigating the use of naltrexone and disulfram in individuals with co-occurring alcohol dependence and bipolar disorder. Recent investigations of psychotherapeutic treatments specifically targeting SUDs in individuals with bipolar disorder also show promise. All individuals with bipolar disorder should be assessed for substance use and counseled concerning the negative impact of substance use and risk for developing dependence.

CONCLUSION

The issue of substance use in those with serious mental disorders is gaining increasing attention from researchers and clinicians with strong evidence that it complicates the natural course of the disorder and is associated with a range of negative outcomes. These negative outcomes relate to a myriad of areas such as general health, psychosocial problems, functional impairment, greater symptomatology, earlier age of onset and treatment. While we have excellent evidence on the epidemiology and prevalence of these co-occurring problems, there remains very little guidance as to evidence-based practice for individuals with more than one disorder.

SEE ALSO

International Data on the Prevalence and Correlates of Comorbid Substance Use and Psychiatric Disorders, Models of Relationships between Substance Use and Mental Disorders, Cannabis Use and the Development and Maintenance of Psychosis, Substance Use and Mood Disorders, Substance Use in Response to Anxiety Disorders

List of Abbreviations

CBT	cognitive behavioral therapy
CRF	corticotrophin releasing factor
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, 4th Edition
ECA	Epidemiologic Catchment Area
NESARC	National Epidemiologic Survey on Alcohol and Related Conditions
NCS	National Comorbidity Study

Glossary

Comorbidity refers to the co-occurrence of an SUD with one or more mental health conditions. The terms "comorbid" and "co-occurring" are used interchangeably throughout this chapter.

Mental disorder refers to the presence of a mental disorder (other than a SUD) as defined by the DSM-IV. Severe mental disorder, for the purposes of this chapter we focus on schizophrenia, major depressive disorder, and bipolar disorder.

Substance use disorder (SUD) refers to the presence of a drug and/or alcohol use disorder as defined by the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV). This includes abuse and dependence.

Further Reading

Baker, A., Velleman, R., 2007. Clinical Handbook of Co-existing Mental Health and Drug and Alcohol Problems. Routledge, London and New York.

Brady, K.T., Sinha, R., 2005. Co-occurring mental and substance use disorders: the neurobiological effects of chronic stress. American Journal of Psychiatry 162, 1483–1493.

Davis, L.L., Uezato, A., Newell, J.M., Frazier, E., 2008. Major depression and comorbid substance use disorders. Current Opinion in Psychiatry 21, 14–18.

de Leon, J., Diaz, F.J., 2005. A meta-analysis of worldwide studies demonstrates an association between schizophrenia and tobacco smoking behaviors. Schizophrenia Research 76, 135–157.

Drake, R.E., Mueser, K.T., Brunette, M.F., McHugo, G.J., 2004. A review of treatments for people with severe mental illnesses and co-occurring substance use disorders. Psychiatric Rehabilitation Journal 27, 360–374.

Grant, B., Stinson, F.S., Hasin, D.S., Dawson, D.A., Chou, S., Ruan, W., et al., 2005. Prevalence, correlates, and comorbidity of bipolar I disorder and axis I and II disorders: results from the national epidemiologic survey on alcohol and related conditions. Journal of Clinical Psychiatry 66, 1205–1215.

Hides, L., Samet, S., Lubman, D.I., 2010. Cognitive behaviour therapy (CBT) for the treatment of co-occurring depression and substance use: current evidence and directions for future research. Drug & Alcohol Review 29, 508–517.

Kavanagh, D.J., Mueser, K.T., 2007. Current evidence on integrated treatment for serious mental disorder and substance misuse. Journal of the Norwegian Psychological Association 44, 618–637.

Large, M., Sharma, S., Compton, M.T., Slade, T., Nielssen, O., 2011. Cannabis use and earlier onset of psychosis: a systematic meta-analysis. Archives of General Psychiatry archgenpsychiatry. 2011.15.

Mueser, K.T., Yarnold, P.R., Levinson, D.F., Singh, H., Bellack, A.S., Kee, K., et al., 1990. Prevalence of substance abuse in schizophrenia: demographic and clinical correlates. Schizophrenia Bulletin 16, 31–56.

Nunes, E.V., Levin, F.R., 2004. Treatment of depression in patients with alcohol or other drug dependence: a meta-analysis. Journal of the American Medical Association 291, 1887–1896.

RachBeisel, J., Scott, J., Dixon, L., 1999. Co-occurring severe mental illness and substance use disorders: a review of recent research. Psychiatric Services 50, 1427–1434.

Schuckit, M.A., 2006. Comorbidity between substance use disorders and psychiatric conditions. Addiction 101, 76–88.

Teesson, M., Hall, W., Proudfoot, H., Degenhardt, L., 2011. Addictions, second ed. Psychology Press, Hove, East Sussex.

CHAPTER

56

Developmental Risk Taking and the Natural History of Alcohol and Drug Use among Youth

Jennifer L. Maggs, Laura Wray-Lake$, John E. Schulenberg***

*The Pennsylvania State University, PA, USA $Claremont Graduate University, CA, USA

**The University of Michigan, MI, USA

OUTLINE

The 2010 Monitoring the Future national survey of high school seniors reported that 71% of adolescents had consumed alcohol by the 12th grade, 42% had tried cigarettes, 44% had used marijuana, and 25% had consumed an illegal drug besides marijuana. These rates suggest that substance use is a common but not universal experience during adolescence. Adolescence is a developmental period characterized by risk-taking behavior, and risk taking is heightened in adolescence for various reasons including biological development and identity exploration.

NORMATIVE DEVELOPMENT OF RISK TAKING

Biology

By late adolescence, most individuals have fully developed in terms of height and reproductive capacity, and by the early 20s, individuals typically reach peak physical functioning in heart and lung strength and athleticism. These physical changes impact adolescents' tolerance of substances, making it possible for youth to recover from negative physical effects, such as hangovers, more quickly. In conjunction with pubertal development, adolescents often look older and desire to be older than they actually are. As alcohol use is a status and privilege of adulthood, the desire to look and act older may prompt initiation of alcohol use.

The adolescent brain offers capacity for heightened sensation seeking and may allow adolescents to derive greater physical pleasure from substance use. It is now recognized that brain development continues beyond adolescence into young adulthood. In particular, higher-order reasoning and cognitive processing capacities continue to develop. Until these capacities fully mature, youth are increasingly likely to engage in excessive substance use and other risk-taking behaviors. Still

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00056-5

developing brain functions can be negatively affected by heavy substance use during adolescence, putting youth at risk for later life repercussions.

Identity

Identity exploration is a fundamental developmental task for adolescents and young adults. This exploration entails recognizing and solidifying one's personal interests, values, goals, commitments, and overall sense of self. As part of identity exploration, adolescents often question assumptions and beliefs that they previously took for granted, including beliefs about substance use. Over the past few decades, US youth have increasingly delayed careers, marriage, and parenthood, and as a result, identity exploration extends beyond adolescence and into young adulthood. These years of prolonged exploration are referred to by some as "emerging adulthood." Although identity exploration is necessary, normal, and healthy, it may sometimes involve experimentation with substance use. Despite the short-term and long-term health risks of the behavior, substance use (and alcohol use in particular) can serve developmental functions such as making friends, testing boundaries of independence, and exploring one's sense of self. Thus, some youth view substance use as a natural pathway to identity formation.

Developmentally Limited Substance Use

Adolescence is a time of rapid development and simultaneous transitions across multiple domains, and this confluence of changes sets the stage for heightened risk-taking behavior. Risk-taking behavior extends to behaviors beyond substance use such as reckless driving, unsafe sex, and extreme sports. Although many youth drink heavily and a sizeable proportion experiment with drugs, only a small number of youth advance to long-term difficulties with substance use including abuse and dependence. As youth grow into adults and take on corresponding responsibilities, experimentation with substance use typically tapers off. Thus, the substance use of most youth is *developmentally limited*. This term refers to substance use that may be heavy during adolescence or young adulthood but is not sustained over time; instead, patterns reflect "aging out" or maturing out of use. Even when youth engage in substance use for a limited period, however, alcohol and drug use can be a highly risky and tragic experience. Moreover, for an important minority of individuals, alcohol and drug use does not decrease. Continued use sets the stage for substance use and abuse problems as well as difficulties in other life domains.

INITIATION OF ALCOHOL AND DRUG USE

The onset of substance use is important to understand because adolescence is a time of greatest risk for the development of substance use disorders. Initiating alcohol use in childhood or early adolescence is a clear marker for problematic patterns of alcohol or drug use in adolescence and adulthood. Age of initiation is defined as the age at which alcohol is first consumed. In measuring alcohol initiation, some researchers exclude sips or tastes, and some only include regular alcohol use. The age cutoff that defines "early" alcohol initiation varies widely across studies. Some researchers define early as use before age 13, whereas others put the cutoff at age 14, 15, or even 18. Regardless of definitional differences, early alcohol initiation consistently predicts later substance use problems. For example, early initiation of alcohol use predicts greater likelihood of drunk driving; more blackouts after use; and increased tolerance, dependence, and substance abuse. Risks increase in correspondence with earlier initiation.

Although a consistent negative link between age of initiation and subsequent alcohol use and disorders has been documented, there is a considerable disagreement over the interpretation of this finding. Some researchers argue that early alcohol use is a cause of later use. In other words, using alcohol in childhood or early adolescence puts individuals on a trajectory toward heightened use and misuse. Prevention efforts consistent with this view seek strategies to delay onset of alcohol use. Taking a different view, other researchers posit that both early and late use stem from the same underlying susceptibility to alcohol problems. Genetic risks for alcohol problems, family environment, or mental illness are some suggested factors that may serve as common causes for initiation and persistence of alcohol use. Proponents of this perspective argue that delaying onset of alcohol use would not reduce problem drinking; rather, prevention efforts should focus on halting the progression from initiation to problem use and abuse. More rigorous research is needed to specify the causal processes involved in the initiation and maintenance of alcohol use in order to determine the most effective prevention efforts.

These two explanations for alcohol initiation are also relevant to understanding the initiation of drug use. Moreover, early alcohol use is a known risk factor for illicit drug use. Middle adolescence is the typical time that individuals start to use marijuana, and late adolescence is a peak time for initiating use of other illicit drugs. Alcohol can serve as a gateway drug, meaning that exposure to alcohol may increase the likelihood that adolescents try other substances and engage in other risk-taking behaviors.

DEVELOPMENTAL TRAJECTORIES

Understanding age-related patterns in alcohol and drug use is necessary from a developmental perspective. A cross-sectional study is a common way to examine substance use at different ages in research; this method compares age groups at a single point in time. While relatively inexpensive to conduct, cross-sectional studies offer incomplete information about how a person's substance use changes as he or she gets older. Long-term longitudinal studies, which follow the same individuals over a period of time, allow researchers to predict individuals' substance use based on their past or present use. These studies chart developmental trajectories, which are patterns that reflect the course of substance use over a number of years. Examining trajectories can help researchers determine why some individuals do not follow patterns that seemed likely based on previous behavior. There are two important types of trajectories: *Average trajectories* demonstrate the overall mean-level developmental pattern for a group of people as a whole, and *multiple trajectories* identify several distinct subgroups of individuals who share the same developmental patterns.

Average Trajectories

Examining average patterns in substance use across adolescence and young adulthood is important for documenting normative developmental changes in use, with normative referring to patterns that describe average change across individuals in a society. Several US longitudinal studies, including Monitoring the Future and the National Longitudinal Survey of Youth, have reported similar average patterns in substance use. Alcohol use increases during adolescence and peaks in the early 20s, at around ages 21–22. This pattern is evident regardless of whether alcohol use is measured as use in the past year, use in the past 30 days, or heavy episodic drinking in the past 2 weeks. Another term for heavy episodic drinking is binge drinking, historically defined in this age group as consuming five or more drinks in a row on a single occasion; binge drinking follows the similar pattern of increasing right after high school, peaking in the early 20s, and then decreasing.

Substance use, and particularly alcohol use, is intertwined with the cultural experience of becoming an adolescent and young adult. Attending college is one such experience that carries high risk for heavy episodic drinking. During high school, adolescents who plan to attend college have *lower* rates of alcohol and other drug use than their peers with no college plans. After high school, *higher* rates of alcohol use and heavy drinking are found among college students compared to their non college-bound peers. However, youth not attending college report higher rates of illicit drug use compared to college attendees.

Studies have also examined trajectories in marijuana use. On average, marijuana use remains relatively stable in the years immediately following high school (ages 18–19) and gradually decreases across the early 20s. Over half of US young adults have reported ever trying marijuana, and 25% have engaged in marijuana use across ages 18–24. However, frequent use is relatively low: In 2008, 5% of US 12th graders reported daily marijuana use, as measured by using marijuana on 20 or more days in the month prior to the survey. Rates of other illegal drugs also tend to peak in late adolescence.

Following the early 20s, substance use steadily declines across young adulthood and beyond. This decline corresponds with the adoption of adult roles. For example, levels of alcohol use, problem drinking, and marijuana use decline in conjunction with marriage, and to a lesser extent with full-time employment and parenthood. Thus, the normative developmental pattern is likely a consequence of shifts in freedoms and responsibilities experienced during adolescence and young adulthood. Adolescence is characterized by freedom and autonomy, whereas the 20s and beyond often involve adult responsibilities. Whether transitions into adult roles cause reduced substance use or vice versa is a matter for debate: Longitudinal research has shown that declines in substance use actually foreshadow marriage, suggesting that the young adults most likely to marry are also those less inclined to use substances.

When examining developmental trajectories, it is important to recognize that substance use is embedded in changing historical contexts. Drugs can change in popularity or availability over time, impacting youth's initiation and use of substances. For instance, nonmedical use of prescription drugs (e.g. OxyContin and Vicodin) and over-the-counter medications (e.g. cough suppressants) have increased in popularity in recent years. As older drugs fall out of favor, new drugs tend to replace them. Alcohol, however, is a mainstay with adolescents. Trends in alcohol use have been relatively stable across several decades, suggesting that risks associated with drinking are prevalent and alcohol use is developmentally relevant regardless of historical time.

The average trajectory approach to examining substance use has several limitations. For example, long periods between measuring substance use in most longitudinal studies can omit important information and lead to overgeneralizations about the simplicity of the overall pattern. Also, as described below, not everyone follows a uniform developmental pathway. Despite these limitations, average developmental trajectories can inform universal prevention efforts, which are aimed at reducing substance use among the population of adolescents. Normative patterns of substance use also

give us valuable insights into youth development and suggest that substance use may serve important developmental functions.

Multiple Trajectories

Rather than identifying a single overall pattern, the multiple trajectory approach focuses on unique subgroups of individuals who follow similar patterns. This approach is useful because it offers the potential for better understanding the origins of substance use disorders and improving early identification of problems. Moreover, individuals follow many different life paths, and substance use patterns likely diverge as a consequence of diverse experiences.

Regarding alcohol use, the most common trajectory subgroup found for adolescents and young adults across several studies is a *low-risk group*. Individuals in this group tend to abstain from drinking, drink small amounts, or consume alcohol rarely across adolescence and young adulthood. Although the exact definition of this group varies across studies, this group engages in very little use and defines one-fifth to two-thirds of adolescents and young adults. In addition, about one-third of adolescents and young adults fall into a *stable-moderate drinker group*. These individuals tend to initiate alcohol use in late adolescence (e.g. age 17 or later); they engage in some heavy drinking over time, but do not increase or decrease use dramatically. These two groups comprise relatively low-risk drinkers and represent the drinking patterns of a large proportion of youth. Thus, the vast majority of youth are not drinking at problematic levels.

Several more problematic patterns in alcohol use have been identified. *Chronic heavy drinkers* tend to initiate heavy drinking early (typically by middle adolescence) and maintain high use in their 20s. *Late onset drinkers* initiate alcohol use later (by middle to late high school), and then steeply increase drinking levels in the late adolescence and young adulthood. Both of these groups are defined by the age at which individuals begin to drink heavily. Because the way researchers define this age varies across studies, these two groups can be difficult to identify or compare across studies.

Another potentially problematic trajectory is exhibited by *fling drinkers*. These individuals tend to engage in heavy drinking during a brief period, typically peaking in adolescence and declining in early adulthood. Fling drinkers make up approximately 10% of adolescents and young adults.

Decreasers reflect a subgroup wherein individuals begin drinking heavily at an early age such as during middle school, looking similar to chronic drinkers. Unlike chronic heavy drinkers who maintain early high levels, this group significantly decreases alcohol consumption during late adolescence and early adulthood. An estimated 10% of adolescents and young adults comprise this group. It is interesting to note that by their mid-20s, however, decreasers not only look different from chronic drinkers in their substance use, but also decreasers are more likely to be married, parents, employed full-time, and financially independent.

Marijuana use trajectories have also been identified in youth aged 18–24. Much like the drinking subgroups, there are two groups of *low-risk marijuana users*: those who abstain and those who report rare and infrequent use across all occasions. *Chronic users*, *fling users*, and *decreasers* are also typical developmental patterns in marijuana use. The vast majority of young people fall into low-risk, decreaser, or fling categories, reflecting either no use or developmentally limited marijuana use. A group of *increasers* display patterns of marijuana use similar to late onset drinkers: these youth report no use or infrequent use in late adolescence which escalates to frequent use in the early 20s. Thus, the multiple trajectory approach has identified very similar alcohol and marijuana use subgroups.

Trajectories in binge drinking and marijuana use across young adulthood vary according to the number of developmental transitions experienced. Typical indicators of young adult developmental transitions include movement into post-secondary education, work, marriage, and parenthood, and out of the parental home. Longitudinal research indicates that experiencing many developmental transitions is associated with a decreasing pattern of substance use; however, experiencing too few developmental transitions is associated with a high, stable developmental pattern of substance use across young adulthood. Although direction of effects was not the focus of the research, findings imply that postponement of developmental transitions is related to higher substance use.

In summary, not everyone experiences the same developmental pathway, and the multiple trajectory approach describes variability in substance use patterns. Identifying unique subgroups based on substance use trajectories can greatly assist researchers and clinicians in pinpointing the most problematic patterns and the best course of prevention or treatment. Both average and multiple trajectory approaches are limited by individuals dropping out of longitudinal studies (called attrition). If attrition is high among the heaviest substance users, then studies will underestimate problems.

RISK AND PROTECTIVE FACTORS

After identifying the developmental patterns in substance use, an important next step is to understand risk and protective factors related to substance use in

adolescence and young adulthood. Risk factors are variables that predict a higher likelihood that a negative outcome will occur, and protective factors refer to anything that reduces or prevents the likelihood of a negative outcome. Risk and protective factors for adolescent alcohol and drug problems have been grouped into at least 17 different categories. Some are distal, contextual factors such as state laws about substance use or ease of obtaining substances in a community. Other risk factors stem from an adolescent's proximal environment, such as experiences in one's family (e.g. conflict with parents), school (e.g. academic failure), and peer group (e.g. rejection). Risk and protective factors can also be person characteristics such as cognition and personality. Elsewhere, scholars have extensively reviewed risk and protective factors; here, we overview a few key domains of developmental significance during adolescence and young adulthood and highlight risk and protective factors found within these domains.

Cognition

A myriad of cognitive changes occur during adolescence that are related to risk for substance use. For example, the capacity to think abstractly and take the perspective of others increases. As adolescents become more independent in their thinking and more likely to question authority, parents', teachers', and other adults' ideas become only one of many possible perspectives. Thus, adults' warnings to avoid alcohol and drugs may be just one of the perspectives that adolescents consider following. In relationships with peers, adolescents become able to view themselves from the perspective of others, and this perspective-taking creates a keen awareness of how actions may impact image and popularity. Adolescents who drink often say they do so to increase their popularity. In fact, with adolescents' increased ability to consider costs and benefits of risky behaviors such as substance use, adolescents become increasingly aware of the potential benefits of substance use and less convinced of its costs. Adolescents' alcohol expectancies, or the expectations that individuals have for positive and negative outcomes of drinking, represent a major risk factor for alcohol use and misuse. When adolescents have more positive attitudes toward alcohol and drug use, they tend to have higher levels of use.

These cognitive risk factors do not imply that adolescents lack cognitive capacity for logical decision making. On the contrary, adolescents' abilities to assess risk and consequences in hypothetical scenarios are equivalent to adults' capabilities by the time adolescents reach age 15. The consensus among researchers is that adolescents use different social and emotional cues when making decisions related to risk behavior. Moreover, certain cognitions serve protective functions: for example, perceived risks to using alcohol or drugs and perceived disapproval of substance use are related to lower likelihood of substance use.

Relationships

Relationships with parents, siblings, peers, and romantic partners can serve as risk or protective factors for substance use. As children become adolescents, they experience more autonomy and increasingly seek independence from parents. As part of this transition, adolescents spend less time with parents and more time in contexts outside of the family. Adolescents tend to increase alcohol use as they become more independent from parents and parents monitor their activities less intensely. Likewise, parental involvement is related to lower substance use for adolescents regardless of gender, ethnicity, or age. Older siblings can also influence substance use. Siblings can be behavioral models by using alcohol or drugs themselves, and they may offer younger siblings access to alcohol or drugs.

Outside of the family, peers represent one of the most important contexts for adolescent development. The majority of youth substance use occurs in the presence of peers, and several different kinds of peer influences can increase risk for alcohol and drug use during adolescence and young adulthood. Unsupervised time spent with friends during adolescence and having friends who get drunk increase the likelihood of substance use. Susceptibility to peer pressure is heightened during transition periods such as transitions to high school and college. At these times in particular, young people are more willing to follow peers' suggestions, which may include using alcohol or other drugs. Adolescents and young adults (particularly college students) tend to overestimate the prevalence of substance use among peers. Perceptions of group norms can encourage heightened drinking as a way to fit in with the peer group. As active agents in their own development, individuals tend to seek out friends with similar interests, values, and behaviors. These similarities among friends can facilitate continuity in behavioral patterns over time. For example, adolescents whose friends use marijuana are more likely to have their own use reinforced in their peer group.

Although less studied, peers can also be positive influences on adolescents, and having peers who exhibit prosocial behaviors is a protective factor. Positive relationships with non-parental adults and greater involvement in social institutions can also serve protective functions against substance use. For example, students who attend church and report high religiosity are less likely to drink heavily or use marijuana across adolescence and young adulthood.

As pubertal development converges with social expectations, adolescents begin to seek romantic and sexual relationships. A range of early sexual experiences are associated with alcohol use during adolescence, and there are several explanations for this co-occurrence. Beliefs in social and sexual enhancement functions of alcohol and other drugs constitutes one possible reason that adolescents are motivated to use substances. In addition, youth may search for social situations where drinking or drugs are present based on a desire to meet new romantic partners. Alcohol consumption can make sexual behaviors more risky by reducing inhibitions and analysis of consequences.

Gender

Gender differences are important for understanding risk and protective factors for substance use. Males have a higher likelihood of belonging to a heavy drinking subgroup and displaying chronic or increasing patterns of marijuana use during adolescence and young adulthood. Adolescent boys are more likely to exhibit externalizing behaviors such as aggression or delinquency, behaviors that are risk factors for substance use. Women are less likely to be in the chronic and late-onset user groups, and are more likely to be represented among abstainers or low-risk users of alcohol and marijuana. Despite level differences, men and women show similar developmental patterns in use. One exception is that women show more dramatic decreases in marijuana use across early adulthood compared to men. Gender differences in substance use seem to develop during middle to late adolescence. At earlier ages, there are generally no gender differences in substance use, and in a few studies, early adolescent females have higher rates of substance use than their male peers.

Race/Ethnicity

White youth tend to be overrepresented in heavy drinking and frequent marijuana use groups compared to ethnic/racial minority youth, and African American and Hispanic American youth are most likely to abstain from substance use. Some evidence suggests that developmental trajectories for binge drinking differ by race/ethnicity, with the pattern of binge drinking for African Americans being flat across ages 18–24, whereas White youth display the overall pattern described above. Overall, however, there are more similarities than differences across ethnic groups in the developmental risk factors and processes related to substance use. Cross-nationally, remarkably similar patterns of alcohol initiation have been found across Western countries, despite wide-ranging laws and adult prevalence rates. This international similarity demonstrates the importance of developmental transitions, which are relatively comparable in Western nations, for understanding substance use.

Achievement

Successful adaptation in educational and occupational roles is a defining feature of healthy development. On the other hand, difficulties in navigating transitions to school and work can exacerbate substance use and related health risks. Therefore, school and work contexts can function as risk or protective factors. Difficulties in transitioning to middle school or high school can put youth at risk for substance use, and truant youth tend to report higher substance use levels across adolescence and young adulthood. On the protective side, high school students with higher GPAs are more likely to abstain from substance use during adolescence and across young adulthood. Positive interactions with family and at school may work together to protect adolescents from health-risk behavior. For example, parental involvement is related to higher school success, which in turn, relates to lower adolescent substance use. Thus, there are important interrelations between protective and risk factors across domains.

In transitioning to college, desires to form social networks and participate in the mythical college party culture may create social motivation to drink heavily. As noted above, college students have higher rates of alcohol use compared to their non-attending peers. In particular, college students are more likely to exhibit patterns of increasing substance use and fling drinking. Stress associated with the academic demands of college may encourage students to turn to heavy drinking or drug use. Living away from home and fraternity or sorority participation represent risk factors for substance use in college; these youth are more likely to display chronic, increasing, or fling patterns. Fraternity members report higher alcohol, marijuana, and other illicit drug use compared to college men not in fraternities. Being in a sorority increases one's risk of using alcohol and marijuana, but does not increase risk of other drug use. Importantly, longitudinal research has revealed that differences between fraternity and sorority members and their peers exist prior to college, suggesting that college-bound youth with a tendency to use substances seek out these experiences through fraternities and sororities. Developmental patterns are also affected: fraternity and sorority members tend to experience greater increases in heavy drinking and marijuana use over time.

Regarding paid work, research findings have consistently shown that working more hours during high school is associated with greater alcohol and other drug use. For example, adolescents who work 16 hours per week or more tend to report more frequent marijuana use. The reason for this finding is controversial,

however. A "third variable" such as disengagement from school may be responsible for causing both increased work hours and substance use. Alternatively, stress from balancing work and school may lead to substance use, or substance use may arise from experiences at work such as exposure to older peers with access to substances. On the protective side, adolescents who set future goals and hold higher occupational aspirations are less likely to engage in substance use.

Developmental Processes of Risk and Protection

A developmental perspective calls for understanding how risk and protective factors interrelate. Some scholars believe that risk and protective factors represent opposite ends of a single continuum. For example, high parental warmth in parent–child relationships can be a protective factor against substance use, whereas low parental warmth can be a risk factor. An alternative view is that protective factors are activated when risk factors are present. In this case, protective factors may counteract effects of risk factors. For example, a supportive family context may be most important in protecting adolescents against substance use only when adolescents are also experiencing negative peer influences. A third possibility is that protective factors function by reducing a risk factor; for example, a supportive family context reduces negative peer influences by improving adolescents' choice of friends.

A developmental perspective on risk and protective factors is needed because these factors change across a person's lifespan and may also change in the way they predict substance use for certain people at certain times. Longitudinal studies make it possible to understand why many individuals do not develop substance use problems despite exposure to documented risk factors. Likewise, a developmental perspective can help explain why individuals without great exposure to risk factors do develop substance abuse or disorders. Multifinality and equifinality are useful concepts in this regard.

Multifinality refers to the idea that any given risk or protective factor can lead to a diversity of outcomes. For example, family history of alcoholism predicts heightened likelihood of alcohol problems in some people, yet reduced likelihood of problems in others. In other words, individuals with the same high-risk background are more likely to become substance use dependent or abstainers. *Equifinality* refers to diverse combinations of risk and protective factors that can result in the same outcome. For example, among heavy-drinking individuals in late adolescence, some individuals may have initiated early due to personality characteristics and antisocial tendencies; other adolescents may exhibit the same behavior, yet experience a short-term heavy drinking episode in response to peer pressures.

Furthermore, timing of risk and protective factors is important for understanding the development of substance use. Early experiences can be useful in predicting substance use and abuse, yet at the same time, some early experiences may be erased or even reversed by later experiences, particularly when a major transition is involved. For example, an individual may have a multitude of early protective factors (e.g. supportive family environment, high self-esteem) that set them on a low-risk path, yet that individual may become a heavy substance user upon transitioning to college and affiliating with substance-using peers. Certain risk and protective factors are stable and robust, meaning that they predict current and future substance use, whereas others are emergent (i.e. predicting future substance use but not current behavior) or concurrent (i.e. only predicting current behavior). In many cases, risk and protective factors that are concurrently related to substance use are not predictive of future behavior.

DEVELOPMENTAL TRANSITIONS

Developmental transitions refer to important changes in individuals' lives that are embedded in cultural contexts. Transitions provide a structure that moves individuals from childhood to adolescence and from adolescence to young adulthood. For example, society holds age-related expectations regarding the timing at which individuals obtain a driver's license or graduate from high school, and societal norms also guide the timing of marriage and parenthood. However, there is considerable diversity in the timing and individuals' experiences of transitions, particularly in the United States. Developmental transitions can vary based on gender, class, and historical period, and individuals can differ in when and how they experience a transition depending on personal goals and life situations. Although all adolescents and young adults do not follow a single developmental pathway, the successful negotiation of transitions in at least a few areas is likely to benefit individuals' health and well-being.

Processes of Transitioning into Substance Use

Transitions help to describe the ways in which alcohol and drug use become a part of developmental experiences. Early initiation and use of alcohol and drugs relates to various negative health and social consequences later in life, yet little is known about what occurs in the interim period between adolescent substance use and later life consequences. Several conceptual models explain how transitions can serve as intervening mechanisms in the development and progression of substance use.

The Overload Model takes the view that alcohol and drug use result from experiencing many developmental transitions over a short period of time. When multiple transitions are experienced simultaneously, stress can accumulate and overwhelm the capacity to cope with the circumstances. In response, individuals may turn to alcohol or drugs as an alternative coping strategy. For example, an adolescent who simultaneously enters college, starts a full-time job, and experiences parental divorce may turn to alcohol or drugs as a way to cope with the stress of multiple transitions.

The Developmental Mismatch Model notes the importance of the fit between individuals and their contexts. For optimal development, individuals' developmental needs should be appropriately matched to opportunities provided by their environments. Some transitions can reduce alcohol and drug use by improving the match between person and context. Other transitions can make the person–environment match less optimal and can adversely affect health. The relationship between work hours and substance use during adolescence supports this model: An adolescent's job may likely be mismatched with the developmental stage in that he or she may be experiencing an increasing need for autonomy yet the job thwarts this goal by not offering opportunities for individual autonomy and being boring or monotonous rather than engaging skills or interests. This mismatch may result in increased substance use.

The Increased Heterogeneity Model argues that difficult transitions intensify individuals' existing strengths and weaknesses. According to this model, individual differences may be exacerbated at critical developmental transitions. During significant transition points, youth who are already experiencing difficulties may have more problems with navigating the new transition, and these difficulties may be reflected in maladjustment, including increases in substance use. Supporting this model, research suggests that individuals who have emotional or academic difficulties tend to experience the transition to college as an immensely stressful event.

The Transition Catalyst Model builds on the idea that a certain degree of risk-taking is a normative developmental experience during adolescence and young adulthood. This model argues that risk-taking behaviors such as alcohol use serve to facilitate healthy development as well as hinder successful transitions. Adolescents may experiment with substance use in order to succeed at certain developmental tasks, such as building bonds with peers, exploring one's identity, or gaining independence from parents. Heavy drinking during the transition to college can help individuals achieve social goals such as new friendships and popularity. At the same time, substance use could turn into a destructive behavior, impede one's developmental goals, or threaten safety and well-being over the short and long term.

In summary, these models, which are examples of some of the more common ones, are important because they reflect the complex and diverse constellation of associations between developmental transitions and substance use patterns. The models are not mutually exclusive, meaning that they could provide overlapping explanations for pathways to substance use and abuse. Furthermore, these models are not the only explanations for the processes by which developmental transitions can facilitate or reduce substance use. Focusing on developmental transitions highlights the role of proximal contexts in providing opportunities as well as constraints for engaging in substance use during adolescence and young adulthood. A developmental perspective recognizes that individuals play an active role in charting their life courses. Individuals choose activities and seek out opportunities based on their personal characteristics, and in these ways, actively shape their own transitions and experiences. In turn, developmental transitions are potential turning points that offer opportunities for increased well-being or risks for negative adaptation.

IMPLICATIONS FOR HEALTH AND PREVENTION

Substance use during adolescence and young adulthood can have harmful effects on health and well-being. Consequences of substance use have biological and social implications, as well as costs to communities and society. These effects can be temporary or long-lasting. Examples of short-term consequences of heavy substance use include an immediate impact on one's judgment and performance, injuries and accidents, and sexual risk-taking. Substance use is likely to lead to long-term negative consequences to the extent that substance use is utilized as a strategy to cope with difficult developmental transitions, use is heavy and sustained over time, or substance use is responsible for a negative life-changing event. Compared to non-heavy drinkers, adolescents who are chronic heavy drinkers have heightened risk for obesity, hypertension, mental illnesses such as depression and anxiety disorders, and substance use and dependence problems in adulthood. These patterns hold even after accounting for important personal characteristics and background factors. Social consequences of heavy substance use include lower likelihood of obtaining a higher education degree and greater criminal behavior such as theft and violence.

Prevention

To counter the numerous serious consequences of substance use during adolescence and young adulthood, a range of prevention efforts are aimed at reducing or eliminating substance use problems. Substance use

prevention efforts are most effective when they are developmentally relevant. Thus, intervention programs should consider the normative developmental change in substance use for the population. A developmental perspective requires taking a long-term view of intervention effects. Changing developmental patterns and the potential for reversal of risk or protective factors emphasize the need to continue assessing individuals who have been involved in a prevention program. It is useful and important to show positive short-term effects of substance use prevention programs, and long-term positive effects are arguably even more important. A prevention program may not demonstrate any short-term improvement, yet long-term effects may accumulate gradually over many years.

The multiple pathways into substance use suggest that a single prevention approach could never fit everyone's unique needs. Instead, multiple approaches to prevention are needed that correspond with unique needs of different subgroups. For instance, chronic heavy drinkers may need intensive early interventions to prevent the escalation of substance abuse and dependence as well as negative health consequences. Fling drinkers may not need intense personalized interventions, and may instead benefit from programs offering high-quality diversions or promoting responsible behaviors. Prevention programs may be successful if focused on strategies to cope with difficult transitions to adulthood, because fling and increasing marijuana users tend to report using marijuana to cope with life difficulties.

Substance use is a leading cause of preventable death among college students. The transition to college comes with multiple developmental tasks including achieving success in academics, romantic relationships, identity, and social life. Given college students' unique needs, high school prevention strategies would not likely translate directly into successful college prevention programs. Programs aimed at easing stresses of the college transition and improving the match between individual expectations and the environment may help youth transition successfully without heightening substance use. Given that alcohol use is heavily embedded into college life, no program is likely to produce sharp reductions in alcohol use, at least in the short term. Thus, harm reduction programs are also important. These programs focus on reducing rather than eradicating heavy drinking behaviors and providing strategies that help young people manage risk behaviors and avoid harmful consequences. Examples of harm-reduction techniques include preventing drunk driving, helping youth know their limits in terms of use, and encouraging accountability among friends at bars or parties. Reducing substance use among fraternity and sorority members is also an important avenue for prevention.

Beyond targeting developmental transitions, prevention programs that target broader social and environmental risk factors have been shown to reduce substance use among adolescents and young adults. For example, changing social norms, heightening penalties for substance use violations, and changing marketing practices can be effective solutions for communities with heavy alcohol use. Overall, despite three decades of prevention research aimed at reducing adolescent substance use, the effectiveness of interventions is mixed. In particular, alcohol use has been stable over historical time and resistant to various local and national efforts to prevent excessive use. The most promising prevention programs operate at multiple levels, targeting, for example, problem-solving skills, peer relationships, social norms, the family environment, and school context. Prevention efforts should also recognize individual and cultural variation in substance use and related risks.

CONCLUSIONS

Adolescence and young adulthood constitute critical transition periods during which initiation and escalation of heavy drinking or drug use can set the stage for problems across the lifespan. Utilizing a developmental perspective entails analyzing risk and protective factors and modeling various functions of developmental transitions. This approach can help us understand initiation of substance use, normative and person-specific developmental patterns, and strategies for preventing destructive patterns of use. Given that risk-taking is a broader phenomenon during adolescence, underlying processes related to the development of substance use are likely applicable to a host of problem behaviors and should also inform optimal development across the lifespan.

SEE ALSO

Adolescent Substance Use: Symptoms and Course, Binge Drinking, Maturing Out, Epidemiology of Adolescent and Young Adult Alcohol, Tobacco, and Drug Use and Misuse in the United States, Cultural Influences on Youth Alcohol and Drug Use.

Glossary

Attrition the proportion of individuals who drop out of later waves in longitudinal studies.

Average trajectory an overall mean-level developmental pattern for a group of people as a whole.

Developmentally limited substance use substance use during a time-limited period of the lifespan, typically adolescence or young adulthood, which is later reduced.

Equifinality diverse combinations of risk and protective factors that can result in the same outcome.

Heavy episodic drinking another term for binge drinking, which has historically been defined in adolescence and young adulthood as consuming five or more drinks in a row on a single occasion.

Multiple trajectories several distinct subgroups of individuals who share the same developmental patterns, typically identified by statistical methods.

Multifinality any given risk or protective factor can lead to a diversity of outcomes.

Further Reading

Arnett, J., 2004. Emerging Adulthood: The Winding Road from the Late Teens through the Twenties. Oxford University Press, New York, NY.

Brown, S., McGue, M., Maggs, J., Schulenberg, J., Hingson, R., Swartzwelder, S., et al., 2008. A developmental perspective on alcohol and youths 16 to 20 years of age. Pediatrics 121 (Suppl. 4), S290–S310. doi:10.1542/peds.2007-2243D.

Chassin, L., Hussong, A., Beltran, I., 2009. Adolescent substance use. Handbook of Adolescent Psychology. In: Individual Bases of Adolescent Development, third ed. vol. 1. John Wiley & Sons Inc, Hoboken, NJ. 723–763.

Johnston, L.D., O'Malley, P.M., Bachman, J.G., Schulenberg, J.E., 2009. Monitoring the Future National Survey Results on Drug Use, 1975–2008. Volume I: Secondary School Students (NIH Publication No. 09-7402). National Institute on Drug Abuse, Bethesda, MD.

Maggs, J., Schulenberg, J., 2004. Trajectories of alcohol use during the transition to adulthood. Alcohol Research and Health 28 (4), 195–201.

Schulenberg, J., Maggs, J., 2002. A developmental perspective on alcohol use and heavy drinking during adolescence and the transition to young adulthood. Journal of Studies on Alcohol (Suppl. 14), S54–S70.

Schulenberg, J., Maggs, J., Hurrelmann, K., 1997. Health Risks and Developmental Transitions during Adolescence. Cambridge University Press, New York, NY.

Schulenberg, J., Merline, A., Johnston, L., O'Malley, P., Bachman, J., Laetz, V., 2005. Trajectories of marijuana use during the transition to adulthood: the big picture based on national panel data. Journal of Drug Issues 35 (2), 255–280.

CHAPTER

57

Epidemiology of Adolescent and Young Adult Alcohol, Tobacco, and Drug Use and Misuse in the United States

Vivian B. Faden, Patricia A. Powell

National Institute on Alcohol Abuse and Alcoholism, Bethesda, MD, USA

OUTLINE

It is now widely recognized that addiction is a developmental disorder with its roots in childhood and adolescence. Adolescents and young adults are among the heaviest users of alcohol, tobacco, and other drugs in the US population. This chapter will provide information on the prevalence of alcohol, tobacco, and other drug use by age and grade, trends in substance use over time, associated consequences, and some thoughts about what all this may mean for the development of America's children. Although the primary focus of this chapter is substance use by adolescents and young adults in the United States, a limited amount of information is also provided about substance use among 15–16-years-olds in 35 European countries and regions based on the 2007 European School Survey Project on Alcohol and Other Drugs (ESPAD).

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00057-7

Published by ELSEVIER Inc.

BACKGROUND

The developmental period encompassing adolescence and young adulthood is one characterized by dramatic changes across multiple biopsychosocial domains. Children are transformed physically, emotionally, and socially as they mature into adults. The extent of change during this transformation is important to keep in mind as one considers the epidemiology of substance use across these life stages, as is the variability in the timing and tempo of these changes across individuals.

Life Stages

Adolescence

The term "adolescence" is often used to describe the period between the onset of puberty and the assumption of adult roles and hence encompasses a broad range of physical, cognitive, and emotional development; grades in school; and levels of independence and responsibility. Not only are adolescents' bodies and brains developing, but they are also experiencing transitions between schools, changes in friends, and peer groups, and adjustments in the way they interact with the world at large. Adolescence is a time of trying new things, as individuals figure out who they are and where they fit in. For many, this includes experimenting with one or more substance. Accordingly, alcohol, tobacco, and drug use typically begin and escalate during this stage of life.

Young Adulthood

Young adulthood, which follows adolescence, is the life stage in which individuals generally settle on a career, choose a life partner, become more financially independent and, for some, start a family. In previous generations, individuals typically married and had children in their late teens or early twenties; over the past 50 years, more and more people are postponing marriage and parenthood until at least their late twenties, and spending their late teens through their mid-twenties in self-focused exploration as they try out different personal relationships and work options. Therefore, in the United States what most people experience during the years from age 18 to 29 has changed dramatically. Essentially, a new developmental stage has been created between adolescence and young adulthood. Jeffrey Arnett has coined the phrase "emerging adulthood" to describe this phenomenon. Unlike in the past, when alcohol and drug use decreased as people took on more adult responsibilities, this relative freedom from responsibility may prolong heavier use of substances throughout a person's twenties.

Substances

This chapter will address the use and abuse of the following substances: alcohol; tobacco; marijuana and hashish; cocaine and specifically crack; heroin; hallucinogens overall and specifically lysergic acid diethylamide (LSD), Phencyclidine (PCP), and ecstasy; inhalants; and nonmedical use of any psychotherapeutic overall, specifically pain relievers (in particular oxycontin), tranquilizers, stimulants (in particular methamphetamine), and sedatives. These substances are considered individually and/or by category depending on the analysis.

Data Sources

This chapter uses data from three US nationally representative surveys conducted in 2009: the National Survey on Drug Use and Health (NSDUH, public use data base), the Monitoring the Future (MTF) survey, and the Youth Risk Behavior Survey (YRBS). Binge drinking is defined in these surveys as drinking five or more drinks on an occasion or in a row.

The NSDUH is an annual, general population, household survey conducted throughout the year by the Substance Abuse and Mental Health Services Administration (SAMHSA) that collects information on individuals 12 years and older. For a number of substances the NSDUH asks questions that parallel the *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edition (DSM-IV) diagnostic criteria for substance abuse and dependence permitting estimation of the prevalence of certain substance use disorders (SUDs) in the general population beginning at age 12.

MTF is a nationally representative school-based survey that has been conducted every spring since 1975 by the University of Michigan Institute for Social Research under a grant from the National Institute on Drug Abuse (NIDA). The study has collected substance use information from 12th graders since its inception and began collecting information from 8th and 10th graders in 1991. In addition to collecting data on substance use, this survey asks respondents about their perceptions of harm and availability of specific substances.

The YRBS is a biennial school-based survey of high school students (9th, 10th, 11th, and 12th graders) conducted by the Centers for Disease Control and Prevention (CDC) since 1991. This survey covers a wide range of risk behaviors, providing a better understanding of how various risk behaviors relate to one another. The Youth Risk Behavior Surveillance System (YRBSS) monitors six categories of health-risk behaviors among youth: (1) behaviors that contribute to

unintentional injuries and violence; (2) tobacco use; (3) alcohol and other drug use; (4) sexual behaviors that contribute to unintended pregnancy and sexually transmitted diseases, including human immunodeficiency virus (HIV) infection; (5) unhealthy dietary behaviors; and (6) physical inactivity. YRBSS includes the national school-based YRBS conducted by CDC and state and local school-based YRBSs conducted by state and local education and health agencies.

EXTENT AND NATURE OF SUBSTANCE USE BY ADOLESCENTS AND YOUNG ADULTS

This section draws on data from the NSDUH, MTF, and the YRBS. NSDUH and MTF data are used for overall prevalence and for information about substance use by basic demographic variables. Past month (or past 30 days in the case of MTF) is emphasized due to the sporadic, or opportunistic, nature of substance use among youth. Some lifetime prevalence data is presented to provide information about the extent to which individuals in various age groups have any experience at all with various substances. Finally, this section draws on data from the YRBS about the relationship of substance use to other risk behaviors.

Lifetime Substance Use among Adolescents and YoungAdults

Table 57.1, based on data from the 2009 NSDUH, presents the lifetime prevalence of substance use by age. This table is useful for understanding the extent to which individuals of various ages have any experience at all

TABLE 57.1 Percentage of Lifetime Substance Use by Age, NSDUH 2009

	Age					
Substance	**12–13**	**14–15**	**16–17**	**18–20**	**21–25**	**26–34**
Tobacco products	9.8	25.0	44.6	63.6	73.6	76.2
Cigarettes	7.2	20.2	37.8	56.7	68.6	71.4
Alcohol any	14.3	37.6	60.7	79.1	90.7	90.6
All illicit drugs	13.6	25.0	40.4	53.8	61.5	59.3
Marijuana/Hashish	3.0	14.6	31.8	47.7	55.9	52.8
Illicit drugs other than marijuana	11.9	17.6	24.1	32.1	40.7	40.6
Cocaine	0.3	0.9	3.6	9.7	18.4	18.4
Crack	0.1	0.1	0.5	1.4	3.9	4.7
Heroin	0.1	0.2	0.5	1.1	2.2	2.1
Hallucinogens	1.2	3.3	7.3	15.1	22.2	23.7
LSD	0.2	0.6	2.0	4.5	8.9	14.5
PCP	0.2	0.4	0.7	0.7	1.9	2.2
Ecstasy	0.6	1.7	4.3	9.5	14.9	15.9
Inhalants	8.1	9.8	9.5	8.9	12.3	13.6
Nonmedical use of psychotherapeutics	5.1	10.5	17.7	25.4	32.1	29.0
Pain relievers	4.3	9.0	15.4	21.5	26.5	22.5
Oxycontin	0.1	1.3	3.2	5.3	7.1	4.5
Tranquilizers	0.6	2.4	5.1	10.7	14.6	12.9
Stimulants	0.8	2.2	3.8	7.8	12.8	12.2
Methamphetamines	0.3	0.7	1.0	2.8	5.7	7.9
Sedatives	0.6	0.7	1.3	1.3	1.8	1.9

Data from the public use data set of the National Survey on Drug Use and Health conducted by the Substance Abuse and Mental Health Services Administration.

with a number of substances. It also provides a background context for considering the 30-day (or monthly) data that follows and which is typically a better indicator of current substance use for youth populations.

An examination of Table 57.1 indicates that 14.3% of US youth ages 12–13 reported having used alcohol, 9.8% reported having used a tobacco product, and 13.6% reported having used an illicit drug; 37.6% of youth ages 14–15 reported having used alcohol, 25.0% reported having used a tobacco product, and 25.0% reported having used an illicit drug; 60.7% of youth ages 16–17 reported having used alcohol, 40.4% reported having used a tobacco product, and 44.6% reported having used an illicit drug, at some point in their lives. Among those ages 18–20, those rates were 79.1, 63.6 and 53.8%. These increasingly high numbers across adolescence speak to the fact that trying substances at some point during this developmental period is normative among youth in the United States.

Prevalence of Substance Use among Adolescents and Young Adults

Table 57.2, based on data from the 2009 NSDUH, presents the prevalence of 30-day use of various substances by age. An examination of Table 57.2 indicates that while adolescents and young adults use many different substances, the most commonly used substances by far are alcohol, tobacco, and marijuana, followed by psychotherapeutics (nonmedical use of pain relievers, tranquilizers, stimulants, and sedatives), hallucinogens,

TABLE 57.2 Percentage of Past-Month Substance Use by Age, NSDUH 2009

	Age					
Substance	**12–13**	**14–15**	**16–17**	**18–20**	**21–25**	**26–34**
Tobacco products	2.4	9.4	21.6	39.2	43.5	39.5
Cigarettes	1.6	7.3	16.8	33.4	37.8	33.7
Alcohol any	3.7	12.4	26.4	49.9	70.4	64.3
Alcohol binge	1.8	6.7	16.9	34.6	46.4	36.3
Alcohol heavy	0.2	1.3	4.5	11.6	15.6	10.2
All illicit drugs	4.0	8.7	16.7	22.5	20.9	12.1
Marijuana/Hashish	0.9	5.9	14.0	20.1	17.6	9.4
Illicit drugs other than marijuana	3.4	4.5	6.2	8.3	8.4	5.1
Cocaine	0.1	0.1	0.5	1.0	1.6	1.0
Crack	0.1	0.0	0.0	0.1	0.2	0.3
Heroin	0.0	0.1	0.1	0.2	0.1	0.1
Hallucinogens	0.2	0.8	1.6	2.5	1.4	0.9
LSD	0.1	0.1	0.2	0.4	0.2	0.0
PCP	0.0	0.1	0.1	0.0	0.0	0.1
Ecstasy	0.1	0.4	0.8	1.5	0.9	0.5
Inhalants	1.6	0.9	1.0	0.6	0.4	0.1
Nonmedical use of psychotherapeutics	1.8	3.4	4.4	6.3	6.4	3.6
Pain relievers	1.5	2.9	3.7	5.1	4.7	2.8
Oxycontin	0.0	0.3	0.6	0.6	0.5	0.4
Tranquilizers	0.2	0.5	1.1	1.7	2.0	1.1
Stimulants	0.2	0.6	0.8	1.2	1.3	0.7
Methamphetamines	0.1	0.2	0.1	0.1	0.3	0.4
Sedatives	0.2	0.1	0.2	0.2	0.3	0.2

Data from the public use data set of the National Survey on Drug Use and Health conducted by the Substance Abuse and Mental Health Services Administration.

and cocaine. Interestingly, the prevalence rates for marijuana and illicit drug use are highest among 18–20-year-olds and decrease substantially among those ages 26–34, from 20.1 to 9.4% for marijuana, and from 22.5 to 12.1% for illicit drugs overall. The prevalence rates for past 30-day tobacco (legal for those 18 and older) and alcohol (legal for those 21 and older) use, by contrast, are highest among 21–25-year-olds at 43.5% for tobacco and 70.4% for alcohol use before dropping off slightly to 39.5 and 64.3% among 26–34-year-olds, respectively.

Importantly, examining the data in Table 57.2 also allows an appreciation of the ramp up in the use of substances across adolescence and into young adulthood. Figure 57.1 presents this dramatic increase graphically for tobacco, alcohol, and marijuana, the substances used by the greatest numbers of adolescents and young adults, and for illicit drugs other than marijuana as a group.

Table 57.3, based on data from the 2009 MTF survey, presents data on the prevalence of past-month use (past 2 weeks for binge drinking) for 8th, 10th, and 12th graders as well as for college students and for non-college students ages 19–28. The substance categories are similar but not identical to those from NSDUH, reflecting differences in the surveys themselves; however, overall the data are remarkably similar. An examination of this table also shows that tobacco, alcohol, and marijuana are the most widely used substances and illustrates the ramp up of substance use with increasing age across adolescence and into young adulthood.

Looking at Substance Use in Adolescence and Young Adulthood by Gender

Table 57.4, based on data from the 2009 NSDUH, presents the prevalence of 30-day use of various substances by sex for various age groups. Interestingly, rates for girls ages 12–13 and 14–15 are higher than or equal to those for boys on most of the measures of substance use reported in the table. On the other hand, once youth are 16–17 years old, males exceed females on all the measures in the table and this is true for all the older age groups as well. One likely explanation is that some younger girls are associating with older boys and are being exposed to substances by them. This notion is supported by studies indicating that early physical maturity in girls is a risk for substance use because of their association with older boys.

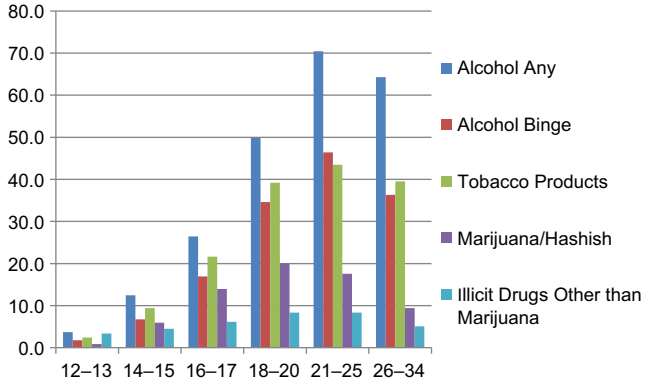


FIGURE 57.1 Substance use increases dramatically during adolescence. *Data from 2009 National Survey on Drug Use and Health conducted by the Substance Use and Mental Health Administration.*

TABLE 57.3 Percentage of Past 30-Day Substance Use by Grade/Age, Monitoring the Future 2009

	Age				
Substance	**8th**	**10th**	**12th**	**College students**	**Young adults 19–28**
Cigarettes	6.5	13.1	20.1	17.9	23.3
Smokeless tobacco	3.7	6.5	8.4	–	–
Alcohol any	14.9	30.4	43.5	65.8	69.4
Alcohol binge*	7.8	17.5	25.2	36.9	36.7
Been Drunk	5.4	15.5	27.4	42.4	40.5
Any illicit drug	8.1	17.8	23.3	20.7	19.8
Marijuana/Hashish	6.5	15.9	20.6	18.5	17
Illicit drugs other than marijuana	3.5	5.7	8.6	8.4	8.5
Cocaine	0.8	0.9	1.3	1.3	1.8
Crack	0.5	0.4	0.6	0.1	0.2
Heroin	0.4	0.4	0.4	0.1	0.2
Hallucinogens	0.9	1.4	1.6	1	0.8
LSD	0.5	0.5	0.5	0.3	0.2
PCP	–	–	0.5	–	**
Ecstasy	0.6	1.3	1.8	0.5	0.6
Inhalants	3.8	2.2	1.2	0.1	0.2
Narcotics other than heroin	–	–	4.1	2.7	3.2
Tranquilizers	1.2	2	2.7	2.2	2.8
Amphetamines	1.9	3.3	3	3.4	2.5
Methamphetamine	0.5	0.6	0.5	0.1	0.3
Sedatives	–	–	2.5	1.2	1.2
Steroids	0.4	0.5	1	0.2	0.3

** indicates 5+ drinks in a row last 2 weeks.*
*** indicates less than 0.05% but greater than 0%.*
– indicates data not available.
Data from Monitoring the Future survey conducted by the University of Michigan under a grant from the National Institute on Drug Abuse.

TABLE 57.4 Percentage of Past-Month Substance Use by Age and Gender, NSDUH 2009

Age	Gender	Tobacco products	Cigarettes	Alcohol any	Alcohol binge	Alcohol heavy	All illicit drugs	Marijuana/ Hashish	Illicit drugs other than marijuana
12–13	Total	2.4	1.6	3.7	1.8	0.2	4.0	0.9	3.4
	Male	2.9	1.5	3.7	1.8	0.2	4.0	0.7	3.4
	Female	1.9	1.7	3.7	1.8	0.2	4.0	1.0	3.4
14–15	Total	9.4	7.3	12.4	6.7	1.3	8.7	5.9	4.5
	Male	10.5	7.1	11.4	6.7	1.3	7.9	6.5	3.2
	Female	8.3	7.5	13.5	6.8	1.3	9.5	5.3	5.9
16–17	Total	21.6	16.8	26.4	16.9	4.5	16.7	14.0	6.2
	Male	25.5	17.8	28.3	18.8	5.1	19.0	16.2	6.7
	Female	17.5	15.8	24.4	14.9	3.8	14.3	11.5	5.6
12–17	Total	11.5	8.9	14.6	8.8	2.1	10.0	7.2	4.7
	Male	13.5	9.2	15.0	9.5	2.3	10.6	8.2	4.5
	Female	9.5	8.6	14.2	8.0	1.8	9.4	6.1	5.0
18–20	Total	39.2	33.4	49.9	34.6	11.6	22.5	20.1	8.3
	Male	45.9	36.8	52.3	39.7	15.6	25.9	23.6	9.0
	Female	31.9	29.7	47.2	29.0	7.2	18.8	16.3	7.6
21–25	Total	43.5	37.8	70.4	46.4	15.6	20.9	17.6	8.4
	Male	52.2	42.8	75.6	56.4	21.9	26.5	22.9	10.2
	Female	35.0	33.0	65.4	36.7	9.4	15.5	12.4	6.6
26–34	Total	39.5	33.7	64.3	36.3	10.2	12.1	9.4	5.1
	Male	46.9	36.7	71.9	46.7	14.4	14.3	11.6	5.6
	Female	32.1	30.7	56.7	26.0	5.9	9.9	7.2	4.6

Data from the public use data set of the National Survey on Drug Use and Health conducted by the Substance Abuse and Mental Health Services Administration.

Looking at Substance Use in Adolescence and Young Adulthood by Race/Ethnicity

Table 57.5, based on data from the 2009 NSDUH, presents the prevalence of 30-day use of various substances by race/ethnicity. At all ages and for most substances, the highest prevalence rates are seen among Whites and American Indians or Alaska Natives. In addition, the biggest differences among race/ethnic groups tend to occur for alcohol and tobacco.

Looking at Substance Use in Adolescence and Young Adulthood by County Type[1]

Table 57.6, based on data from the 2009 NSDUH, presents the prevalence of 30-day use of various substances

[1]Counties were grouped based on the "Rural/Urban Continuum Codes" developed by the US Department of Agriculture (2003). Each county is in either a metropolitan statistical area (MSA) or outside of an MSA (also see Butler and Beale, 1994). Large metropolitan (large metro) areas have a population of 1 million or more. Small metropolitan (small metro) areas have a population of fewer than 1 million. Nonmetropolitan (nonmetro) areas are outside of MSAs and include urbanized counties with a population of 20 000 or more in urbanized areas, less urbanized counties with a population of at least 2500 but fewer than 20 000 in urbanized areas, and completely rural counties with a population of fewer than 2500 in urbanized areas. Estimates based on county-type information presented in this report use the 2003 revised definition of an MSA; estimates for 2002 in this report, therefore, are not directly comparable with those presented in the 2002 NSDUH report (Office of Applied Studies (OAS), 2003).

TABLE 57.5 Percentage of Past-Month Substance Use by Age and Race/ethnicity, NSDUH 2009

Age	Race/ethnicity	Tobacco products	Cigarettes	Alcohol any	Alcohol binge	Alcohol heavy	All illicit drugs	Marijuana/ Hashish	Illicit drugs other than marijuana
12–17	Total	11.5	8.9	14.6	8.8	2.1	10.0	7.2	4.7
	Hispanic or Latino	9.0	7.3	15.4	9.7	2.2	10.9	7.8	4.6
	Not Hispanic or Latino								
	White	13.8	10.6	16.0	9.9	2.6	9.7	7.1	4.6
	Black or African American	7.4	5.2	10.2	4.8	0.5	10.9	7.1	5.4
	American Indian or Alaska Native	17.4	12.0	11.3	9.7	1.7	14.1	10.7	5.5
	Native Hawaiian or other Pacific Islander	12.7	12.5	25.8	14.4	0.1	12.9	11.0	4.3
	Asian	3.2	2.3	6.0	2.2	0.3	5.7	2.5	4.0
	Two or more races	15.2	13.6	16.9	10.8	1.7	12.6	10.5	5.3
18–20	Total	39.2	33.4	49.9	34.6	11.6	22.5	20.1	8.3
	Hispanic or Latino	30.4	27.1	44.2	30.0	6.2	18.2	14.8	7.2
	Not Hispanic or Latino								
	White	46.0	38.9	55.7	40.6	15.5	24.9	22.4	10.1
	Black or African American	26.8	22.8	38.4	19.8	4.1	20.2	18.9	4.3
	American Indian or Alaska Native	58.7	49.5	36.9	26.9	10.5	26.5	25.7	4.1
	Native Hawaiian or other Pacific Islander	49.2	44.7	41.4	33.2	13.0	26.1	26.1	10.0
	Asian	16.9	14.3	32.1	20.6	3.4	10.6	8.5	3.0
	Two or more races	44.9	35.3	53.5	37.5	16.5	28.9	28.5	8.3
21–25	Total	43.5	37.8	70.4	46.4	15.6	20.9	17.6	8.4
	Hispanic or Latino	33.1	30.8	56.7	38.1	9.4	14.7	10.6	6.4
	Not Hispanic or Latino								
	White	49.6	42.6	77.4	53.6	20.0	23.5	20.0	10.0
	Black or African American	34.9	29.1	61.5	32.4	8.0	20.5	18.3	4.8
	American Indian or Alaska Native	48.4	45.8	56.6	41.8	15.5	33.5	20.9	21.5
	Native Hawaiian or other Pacific Islander	50.0	48.2	70.3	40.7	21.9	21.1	19.3	9.0
	Asian	25.6	24.3	59.2	27.6	5.1	8.9	7.1	3.3
	Two or more races	56.6	46.0	80.9	49.9	9.2	31.0	28.4	7.9
26–34	Total	39.5	33.7	64.3	36.3	10.2	12.1	9.4	5.1
	Hispanic or Latino	33.6	29.7	54.5	35.6	7.9	9.0	7.0	3.7

(*Continued*)

TABLE 57.5 Percentage of Past-Month Substance Use by Age and Race/ethnicity, NSDUH 2009—cont'd

Age	Race/ethnicity	Tobacco products	Cigarettes	Alcohol any	Alcohol binge	Alcohol heavy	All illicit drugs	Marijuana/ Hashish	Illicit drugs other than marijuana
	Not Hispanic or Latino								
	White	44.5	37.9	69.3	39.9	12.3	13.9	10.7	6.3
	Black or African American	34.5	28.4	64.1	30.0	7.3	10.4	9.0	2.7
	American Indian or Alaska Native	51.4	48.8	43.1	30.8	10.3	20.4	14.2	7.6
	Native Hawaiian or other Pacific Islander	30.6	15.3	57.1	25.2	1.8	1.4	1.4	1.1
	Asian	15.9	14.4	47.1	15.3	1.6	6.6	4.5	3.1
	Two or more races	51.2	43.6	75.9	42.2	16.3	16.1	15.3	4.4

Data from the public use data set of the National Survey on Drug Use and Health conducted by the Substance Abuse and Mental Health Services Administration.

TABLE 57.6 Percentage of Past-month Substance Use by Age and County Type, NSDUH 2009

Age	Metro	Tobacco products	Cigarettes	Alcohol any	Alcohol binge	Alcohol heavy	All illicit drugs	Marijuana/ Hashish	Illicit drugs other than marijuana	Cocaine	Crack	Methamphet-amines
12–17	Total	11.5	8.9	14.6	8.8	2.1	10.0	7.2	4.7	0.2	0.0	0.1
	Nonmetro	14.5	10.5	14.8	9.4	2.3	8.5	5.3	4.7	0.2	0.1	0.2
	Small metro	12.0	9.6	14.7	8.9	2.3	10.5	7.8	5.1	0.3	0.0	0.1
	Large metro	10.4	8.0	14.6	8.5	1.9	10.2	7.3	4.5	0.2	0.0	0.1
18–20	Total	39.2	33.4	49.9	34.6	11.6	22.5	20.1	8.3	1.0	0.1	0.1
	Nonmetro	46.4	39.1	47.8	35.8	11.5	18.6	15.2	8.0	1.0	0.0	0.3
	Small metro	40.6	33.8	52.2	36.6	12.7	24.0	20.9	9.3	1.0	0.1	0.0
	Large metro	35.8	31.2	49.1	32.9	10.9	22.9	21.2	7.8	1.1	0.0	0.1
21–25	Total	43.5	37.8	70.4	46.4	15.6	20.9	17.6	8.4	1.6	0.2	0.3
	Nonmetro	52.3	45.6	69.6	45.5	14.9	16.6	13.3	7.6	0.6	0.1	0.6
	Small metro	44.3	38.8	69.9	47.3	17.5	21.3	17.4	9.5	2.1	0.2	0.3
	Large metro	40.6	35.1	70.9	46.1	14.7	21.9	18.8	7.9	1.5	0.1	0.2
26–34	Total	39.5	33.7	64.3	36.3	10.2	12.1	9.4	5.1	1.0	0.3	0.4
	Nonmetro	51.9	45.0	58.3	36.3	7.9	11.3	7.6	6.2	0.4	0.1	0.1
	Small metro	40.5	35.1	63.0	34.9	10.6	11.2	8.8	4.7	1.0	0.5	0.3
	Large metro	35.7	30.1	66.5	37.1	10.5	12.8	10.2	5.0	1.2	0.2	0.5

Data from the public use data set of the National Survey on Drug Use and Health conducted by the Substance Abuse and Mental Health Services Administration.

by county type. While there is some variation in substance use across county type (nonmetropolitan, small metropolitan, and large metropolitan) and age group, no clear generalizations can be made.

Long-Term Trends in Adolescent Substance Use

MTF which has surveyed 12th grade students in the United States since 1975 provides an excellent opportunity to look at long-term trends in 12th grade substance use.

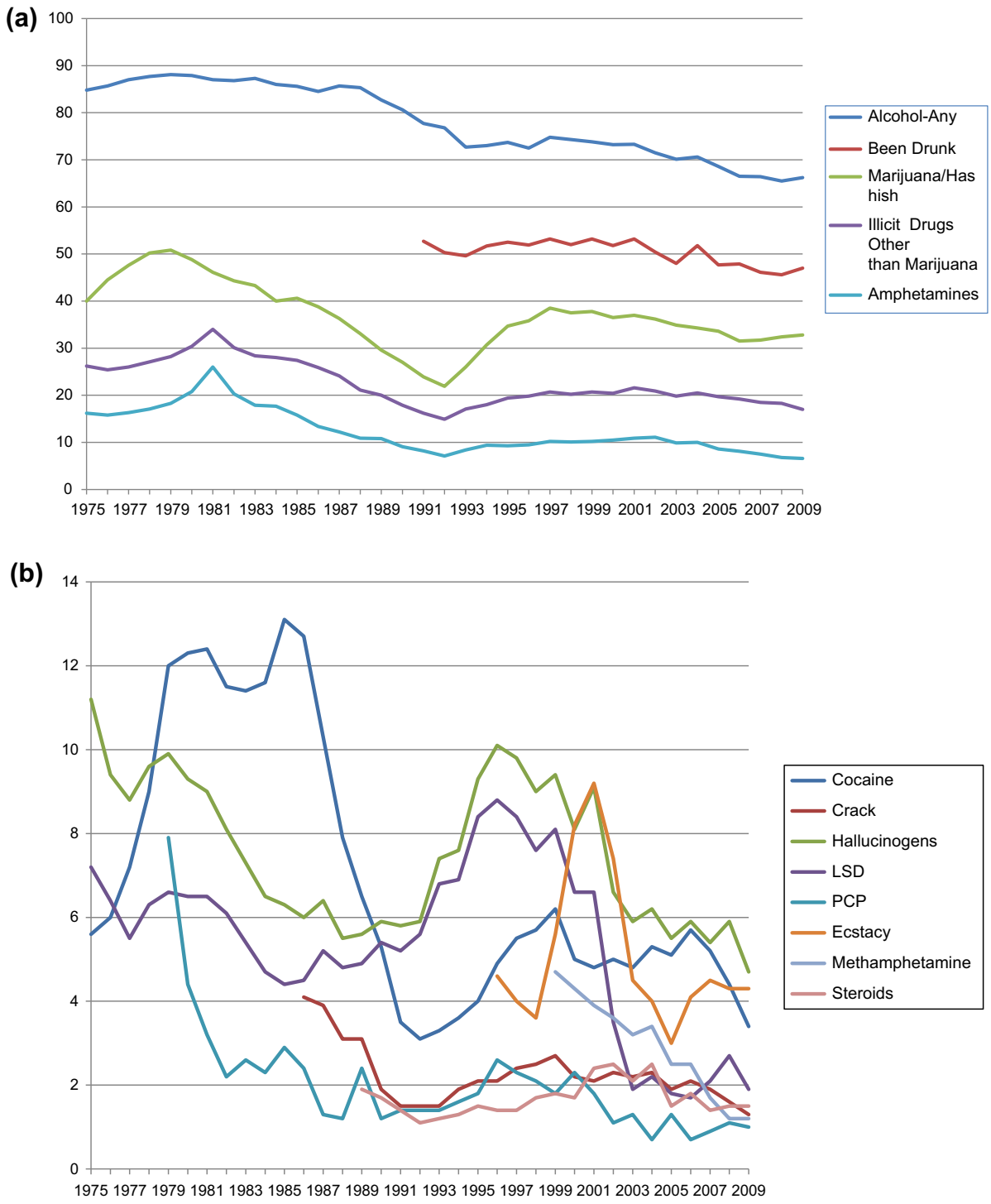


FIGURE 57.2 (a) Trends in annual substance use by 12th graders 1975–2009 – Higher prevalence substances. (b) Trends in annual substance use by 12th graders 1975–2009 – Lower prevalence substances. *Data from Monitoring the Future survey conducted by the University of Michigan under a grant from the National Institute on Drug Abuse.*

Figures 57.2a and b present data illustrating how the prevalence rates of use of various substances have changed across the period from 1975 until 2009. An examination of Fig. 57.2a, which illustrates trends in those substances that are used by more young people, shows that more youth have experience with alcohol and being drunk followed by the use of marijuana/hashish, illicit drugs other than marijuana, and amphetamines. This ranking of use

has remained consistent across the entire period for which data were collected. Of note, use of all substances in this figure declined between 1985 and 1992. Whereas the prevalence of marijuana use by 12th graders increased substantially between 1992 and 1997 before leveling off and then declining slightly, the prevalence of use of the other substances in this figure remained relatively constant or declined somewhat from 1992 to 2009.

Figure 57.2b depicts similar information for less widely used substances and allows an appreciation of the fads in drug use among young people over time. In fact, for some of the substances shown, data collection does not span the entire time period but rather was initiated as a result of the increased popularity of these substances. For other substances, such as ecstasy and cocaine, the figure captures both the steep rise and fall in their use by 12th graders. The figure also illustrates that different substances were popular at different times; for example, cocaine use peaked in 1985 whereas ecstasy use peaked in 2001.

Relationship of Substance Use to Other Risk Behaviors

Data on risk behaviors from the 2009 YRBS allow an examination of the concurrent use of various substances as well as the relationship between substance use and other risk behaviors. The data from 9th, 10th, 11th, and 12th graders were collapsed to provide an overview for purposes of this chapter. Table 57.7a presents the relationship among use of five substances: alcohol (binge drinking one or more times in the past 30 days), smoking (smoked cigarettes on 20 or more of the past 30 days), marijuana (used or smoked one or more times in the past 30 days), cocaine (ever used any form), and methamphetamines (ever used). Not surprisingly, there is a strong association between use of the five substances. Binge drinking is strongly associated with marijuana use and vice versa; and both binge drinking and marijuana use are similarly associated with regular smoking (20.9 and 25.7% respectively), with cocaine use (18.0 and 21.4% respectively) and with methamphetamine use (10.8 and 13.2% respectively). Among those who ever used cocaine or methamphetamine almost 75% have also used marijuana, about 70% binge drink, and 40% smoke regularly.

Table 57.7b presents the relationship of the use of the same five substances with a number of non-substance high-risk behaviors. In general, of the five substances analyzed, ever having used cocaine or ever having used methamphetamine are more strongly associated with the non-substance high risk behaviors in the table than are binge drinking, regular smoking or marijuana use.

Attitudes about Substance Use among Adolescents

MTF collects information from 8th, 10th, and 12th graders about perceived risk of using various substances. Importantly, trends in use often reflect changing attitudes. For example, Fig. 57.3a shows the dramatic inverse relationship between 12th graders' attitudes about perceived harm from marijuana use and the prevalence of its use between 1975 and 2009.

As perceived risk increases, a decline in use often follows, and similarly as perceived risk decreases, a rise in use often follows, albeit sometimes with a slight time lag. Figure 57.3b presents perceived risk among 12th graders over time (early 1990s to 2009) for some of the most frequently used substances. For 8th and 10th graders (data not presented) use of all substances is generally lower and perceived risk is generally higher than for 12th graders, but trends in attitudes and use over time are similar. Perceived risk for alcohol and

TABLE 57.7A Prevalence of Other Substance Use (Row Percentage) among Individuals with a Given Substance Use, YRBS, 2009

	Five+ drinks 1+ past 30 days	Used marijuana 1+ times past 30 days	Used cocaine 1+ times in life	Used meth 1+ times in life	Smoked 1+ past 30 days	Smoked on 20+ past 30 days
Five+ drinks 1+ past 30 days	100.0	51.6	18.0	10.8	50.8	20.9
Smoked on 20+ past 30 days	69.2	71.6	34.2	20.1	100.0	100.0
Used marijuana 1+ times past 30 days	60.9	100.0	21.4	13.2	56.8	25.7
Used cocaine 1+ times in life	70.6	72.3	100.0	42.2	74.6	41.5
Used meth 1+ times in life	68.8	73.4	70.3	100.0	72.9	40.6

Data from the Youth Risk Behavior Surveillance System conducted by the Centers for Disease Control and Prevention.

TABLE 57.7B Prevalence of Other Risk Behaviors (Row Percentage) among Individuals with a Given Risk Behavior, YRBS 2009

	Rode 1+ times with drinking driver past 30 days	Carried weapon 1+ times past 30 days	Carried gun 1+ past 30 days	Had sex with 4+ people in life	Of current sex, used condom last time	Of current sex, used alcohol or drugs last time	Made suicide plan 12 months
Five+ drinks 1+ past 30 days	56.3	30.6	12.1	29.4	60.7	36.6	15.2
Smoked on 20+ past 30 days	57.6	38.5	16.2	45.2	50.6	42.4	21.5
Used marijuana 1+ times past 30 days	52.3	31.6	12.5	33.5	57.3	42.1	16.5
Used cocaine 1+ times in life	64.8	45.4	22.9	46.8	46.3	51.4	28.9
Used meth 1+ times in life	64.7	54.6	30.6	53.0	46.1	56.1	36.7

Data from the Youth Risk Behavior Surveillance System conducted by the Centers for Disease Control and Prevention.

tobacco use have been stable or increased slightly over the past few years while perceived risk for illicit drug use has declined. Correspondingly, alcohol and tobacco use have decreased slightly while the use of a number of illicit substances has increased somewhat in recent years.

Information about the perceived availability of a substance (how easy they think it would be to get that substance) by youth is also collected by MTF. As was seen with attitudes, trends in use of different substances track perceived availability. This may in part reflect increased opportunity to obtain various substances from greater numbers of friends who are using them.

CONSEQUENCES OF UNDERAGE AND YOUNG ADULT SUBSTANCE USE

The consequences of underage and young adult substance use are many and varied. Depending on the substance, they include social, academic, and vocational problems; unwanted, unintended, and/or unprotected sexual activity; unintentional injuries including from car crashes and falls; involvement with the criminal justice system; truancy; the potential for possible irreversible, adverse effects on the developing brain; increased risk for HIV/AIDS; compromised present and future health; the development of SUDs as a teen or young adult; increased risk for substance dependence in later life; and deaths from a substance overdose. Some of the more readily available prevalence estimates of consequences are highlighted below.

It is estimated that each year approximately 5000 people in the United States under the age of 21 die as a result of drinking alcohol. This includes about 1900 deaths from motor vehicle crashes, 1600 as a result of homicides, 300 from suicides, as well as hundreds from other injuries such as falls, burns, and drownings. In addition, early initiation of alcohol use is associated with increased risk for developing alcohol dependence at some point later in life, as well as developing dependence more quickly and at younger ages.

Many sources including the NSDUH and MTF have documented high levels of drinking and binge drinking among college students. Thus, it is not surprising that there are a high number of negative alcohol-related consequences that occur among college students including date rapes, assaults, and unintentional injuries. It is estimated that about 1800 alcohol-related deaths among college students ages 18–24 occur each year (*see* Costs and Consequences (Morbidity and Mortality) Associated with Adolescent and College Drinking and Related Problems).

Use of other substances clearly poses risks as well. For example, the use of tobacco by underage and young adult individuals increases the likelihood of nicotine addiction and the devastating physical illnesses that results from chronic tobacco use, such as lung cancer and emphysema.

Use of illicit drugs and nonmedical use of prescription drugs can result in drug overdose. According to the CDC, in 2007 about 93% of the 29 846 unintentional poisoning deaths that occurred in the United States were unintentional drug (illegal, prescription or over-the-counter drug) overdoses. Using this overall percentage to estimate the number of drug overdoses among 12–17, 18–20, 21–25, and 26–34-year-olds from the reported number of unintentional poisonings in these age groups translates to approximately 300, 891, 2322, and 4751 drug overdoses in each age group, respectively.

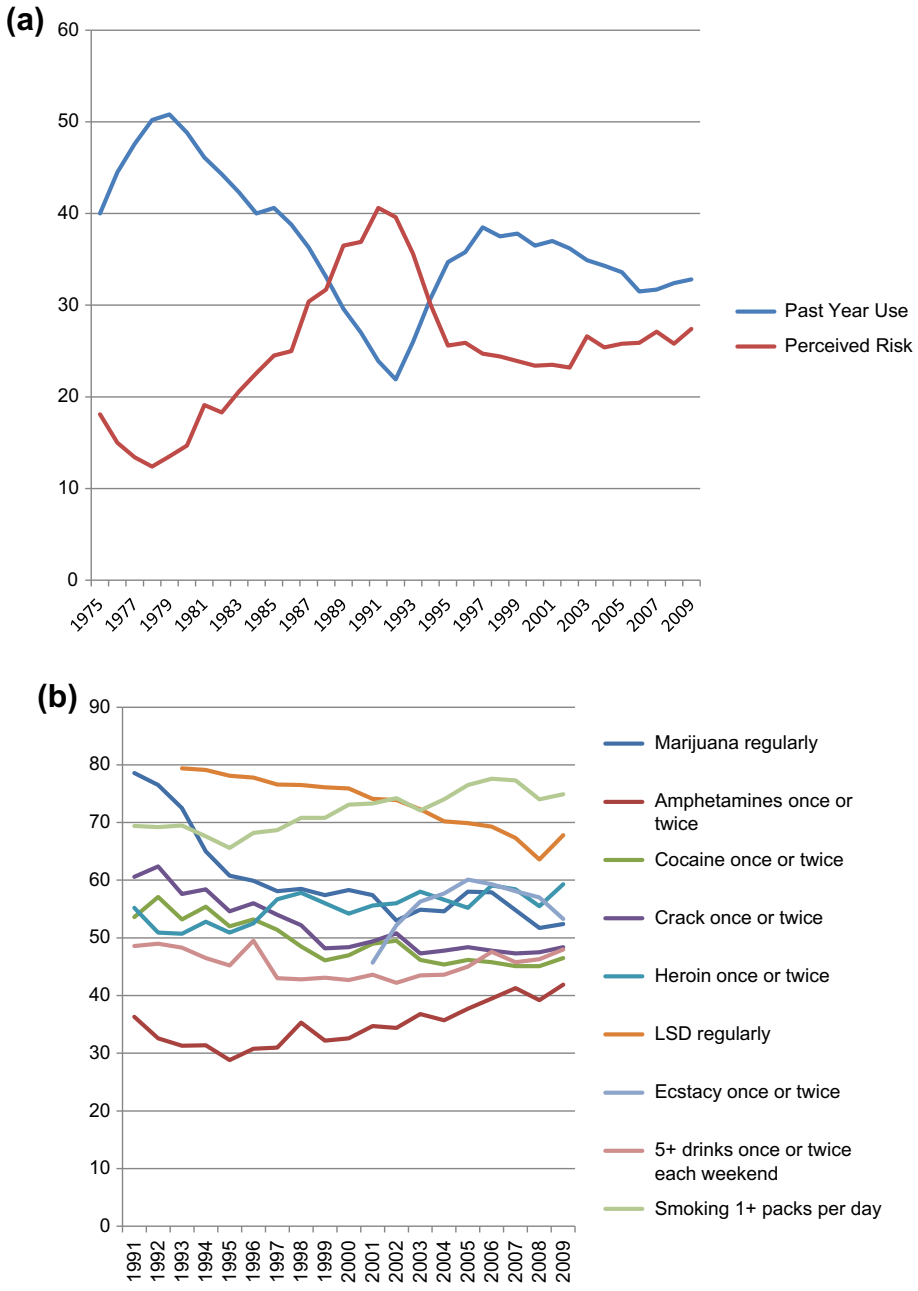


FIGURE 57.3 (a) 12th graders' past-year marijuana use versus perceived risk of occasional marijuana use, 1975–2009. (b) 12th graders' perceived risk of use of various substances 1991–2009. *Data from Monitoring the Future survey conducted by the University of Michigan under a grant from the National Institute on Drug Abuse.*

EXTENT AND NATURE OF SUDs AMONG ADOLESCENTS AND YOUNG ADULTS

The development of an SUD is an important risk of substance use by adolescents and young adults. The sections below discuss the epidemiology of SUDs in this population, their comorbidity with other psychiatric disorders, and the receipt of treatment by those who develop an SUD.

The Prevalence of Substance Abuse Disorders in Adolescence and Young Adulthood

Table 57.8a, based on data from the 2009 NSDUH, presents the prevalence of SUDs among adolescents and young adults in the general US population by age and substance. An examination of this table reveals that more individuals are addicted to nicotine (by far the highest number) and to alcohol, than to other substances.

TABLE 57.8A Percentage of Past-Year Abuse and Dependence by Substance and Age, NSDUH 2009

Abuse and dependence	Substance	12–13	14–15	16–17	18–20	21–25	26–34
Abuse	Illicit drugs	0.6	1.9	3.4	3.0	1.9	1.2
	Marijuana/Hashish	0.2	1.4	2.9	2.6	1.3	0.8
	Nonmedical use of psychotherapeutics	0.3	0.5	0.8	0.6	0.5	0.3
	Alcohol	0.4	2.5	4.7	9.1	9.9	5.8
	Both illicit drugs and alcohol	0.1	0.3	1.0	1.0	0.6	0.4
	Illicit drugs or alcohol	0.9	4.1	7.2	11.1	11.2	6.5
Dependence	Illicit drugs	0.6	1.8	4.4	6.1	5.1	2.8
	Marijuana/Hashish	0.3	1.4	3.6	4.6	3.2	1.4
	Nonmedical use of psychotherapeutics	0.2	0.5	0.7	1.7	1.6	1.0
	Alcohol	0.3	1.7	3.6	5.6	7.2	5.9
	Both illicit drugs and alcohol	0.1	0.4	1.0	1.6	1.4	0.8
	Illicit drugs or alcohol	0.8	3.2	6.9	10.1	10.9	7.9
	Nicotine (past month)—NDSS	0.3	1.3	3.7	9.6	12.7	12.9
	Nicotine (past month)—FTND	0.5	1.9	4.5	10.7	13.3	13.2
	Nicotine (past month)—either NDSS or FTND	0.5	2.3	5.8	13.8	17.0	17.0
Abuse or dependence	Illicit drugs	1.1	3.7	7.8	9.1	6.9	3.9
	Marijuana/Hashish	0.5	2.8	6.6	7.2	4.5	2.1
	Nonmedical use of psychotherapeutics	0.5	1.0	1.5	2.3	2.1	1.3
	Alcohol	0.7	4.3	8.3	14.7	17.1	11.7
	Both illicit drugs and alcohol	0.4	1.5	3.5	4.5	3.4	1.9
	Illicit drugs or alcohol	1.4	6.5	12.6	19.3	20.6	13.7

Data from the public use data set of the National Survey on Drug Use and Health conducted by the Substance Abuse and Mental Health Services Administration.

Table 57.8b, based on data from the 2009 NSDUH, presents the percentage of current users of a substance who meet the criteria for a diagnosis of abuse or dependence for that same substance which likely reflects the addiction liability of that substance. This table indicates that addiction liability (likelihood of having a SUD given that one uses the substance) is highest for nicotine for all age groups. There is some variability by age, but in general the next highest percentages are for illicit drugs, followed by marijuana, psychotherapeutic drugs used nonmedically, and then alcohol.

These tables considered together indicate that addiction risk is highest for nicotine and quite high for illicit substances, and that alcohol and nicotine affect the largest numbers of people.

Comorbidity of SUDs and Other Psychiatric Conditions in Underage and Young Adult Populations

Understanding the relationship of SUDs and other psychiatric disorders, including other SUDs, is very important as those youth/young adults who have more than one disorder have more complex prevention and treatment needs.

TABLE 57.8B Percentage of Past-Year Abuse and Dependence by Substance and Age, among Respective Substance Users, NSDUH 2009

Abuse and dependence	Substance	12–13	14–15	16–17	18–20	21–25	26–34
Abuse	Illicit drugs	6.6	10.1	11.2	7.7	5.3	5.4
	Marijuana/Hashish	7.7	11.3	11.8	7.7	4.4	4.7
	Nonmedical use of psychotherapeutics	7.7	6.6	6.6	3.6	3.5	3.6
	Alcohol	4.2	8.6	9.4	12.8	11.7	7.2
	Both illicit drugs and alcohol	1.8	2.5	3.9	2.6	1.7	1.8
	Illicit drugs or alcohol	6.1	11.6	13.0	15.1	13.1	8.1
Dependence	Illicit drugs	6.7	9.9	14.2	15.8	14.5	12.9
	Marijuana/Hashish	13.0	11.3	14.6	13.4	11.0	8.3
	Nonmedical use of psychotherapeutics	7.1	7.1	6.0	11.2	11.1	10.7
	Alcohol	3.6	5.9	7.1	7.9	8.6	7.4
	Both illicit drugs and alcohol	4.3	3.3	3.9	4.4	4.3	3.9
	Illicit drugs or alcohol	5.2	9.0	12.6	13.7	12.7	9.7
	Nicotine (past month)—NDSS	18.1	17.3	22.1	28.8	33.7	38.1
	Nicotine (past month)—FTND	28.9	26.0	27.0	32.0	35.1	39.2
	Nicotine (past month)—either NDSS or FTND	34.4	31.5	34.4	41.3	45.0	50.4
Abuse or dependence	Illicit drugs	13.2	20.0	25.5	23.4	19.8	18.3
	Marijuana/Hashish	20.7	22.6	26.4	21.1	15.5	13.0
	Nonmedical use of psychotherapeutics	14.9	13.6	12.6	14.7	14.5	14.3
	Alcohol	7.7	14.5	16.5	20.7	20.3	14.6
	Both illicit drugs and alcohol	13.2	11.8	13.7	12.3	10.1	9.1
	Illicit drugs or alcohol	9.7	18.4	22.8	26.3	24.1	16.9

Data from the public use data set of the National Survey on Drug Use and Health conducted by the Substance Abuse and Mental Health Services Administration.

As is seen among adults, for certain youth substance use problems are related to psychiatric comorbidity. These can include, for example, depression, anxiety, conduct disorder, and various personality disorders. The NSDUH provides information on the co-occurrence of major depression, both lifetime and past year, and dependence on various substances. Table 57.9, based on data from the 2009 NSDUH, illustrates that in the adolescent and young adult population, many who have an SUD also have experienced a major depressive episode within the past year; the data indicate that the prevalence of a lifetime depressive episode in this population is even higher. Table 57.9 also shows that among adolescents and young adults, just as the use of one substance is associated with the use of other substances (see Table 57.7), so too is dependence on one substance associated with dependence on other substances.

Receipt of Treatment for Substance Use Problems by Adolescents and Young Adults

In spite of the high prevalence of alcohol and other substance use problems among underage and young adult populations in the United States (as determined in nationally representative surveys), only a small proportion of those with such a disorder receive treatment. Table

TABLE 57.9 Percentage of Major Depressive Episode and Substance Use Dependence, Age, NSDUH 2009

	Major depressive episode		Dependence			
Age	**Lifetime**	**Past year**	**Illicit drugs**	**Marijuana/Hashish**	**Nonmedical use of psychotherapeutics**	**Alcohol**
Dependence—Alcohol						
12–17	36.8	27.3	27.4	19.1	10.2	100.0
18–20	29.4	20.9	29.0	20.3	10.1	100.0
21–25	28.4	21.9	20.0	12.4	7.5	100.0
26–34	29.2	20.6	13.4	5.9	4.1	100.0
Dependence—Marijuana/Hashish						
12–17	29.1	21.9	100.0	100.0	8.4	20.3
18–20	24.0	15.6	100.0	100.0	13.3	24.8
21–25	29.1	17.6	100.0	100.0	8.9	27.7
26–34	31.4	14.2	100.0	100.0	5.2	25.5
Dependence—Nonmedical use of Psychotherapeutics						
12–17	61.6	52.5	100.0	30.5	100.0	39.4
18–20	33.9	27.5	100.0	34.7	100.0	32.4
21–25	43.6	34.3	100.0	17.9	100.0	33.4
26–34	43.9	28.5	100.0	7.3	100.0	25.0
Dependence—Illicit Drugs						
12–17	35.1	27.6	100.0	78.8	21.7	23.0
18–20	25.4	17.6	100.0	74.5	28.5	26.4
21–25	32.7	22.6	100.0	64.0	31.9	28.5
26–34	33.2	19.1	100.0	49.4	35.3	28.8

Data from the public use data set of the National Survey on Drug Use and Health conducted by the Substance Abuse and Mental Health Services Administration.

57.10, based on data from the 2009 NSDUH, illustrates the receipt of treatment in the past year and lifetime treatment for those individuals who had past year abuse or dependence. It clearly shows that, at all ages, just a fraction of those who need treatment actually receive it, even among dependent individuals.

Beyond the United States – Youth Substance Use in Europe

The information presented in this section is from the 2007 ESPAD which collects data from 15 to 16-year-old students in 35 European countries and regions.[2] Given the ESPAD survey's similarity to MTF, comparisons can be made between youth substance use in the United States and Europe. For comparability, data for the United States in this section is from the 2007 MTF.

For past 30-day alcohol use as well as for past 30-day cigarette use among 15-to 16-year-old students, the United States is in the bottom three countries along with Iceland and Armenia. With respect to having been drunk, youth in Denmark report by far the highest rates. Students in the United Kingdom, Austria, Ireland, the Slovak Republic, and Spain report the next highest levels. The middle group of countries includes the United States and many of the Eastern European and Northern European nations as well as France. The lowest rates of having been drunk are reported by students in many of the Southern European countries as well as Sweden, the Netherlands, and others. With respect to illicit drugs, the United States tops the list in

[2]Data from Spain is from a separate national survey.

TABLE 57.10 Percentage Receiving Treatment by Substance for All AUD and SUD Individuals Ages 12–34, NSDUH 2009

	Past-year abuse				Past-year dependence			
Substance treatment	**Illicit drugs**	**Alcohol**	**Both illicit drugs and alcohol**	**Illicit drugs or alcohol**	**Illicit drugs**	**Alcohol**	**Both illicit drugs and alcohol**	**Illicit drugs or alcohol**
Age 12–13								
Lifetime treatment for drug or alcohol use	4.6	8.6	0.0	6.6	17.6	8.2	19.1	13.3
Past-year treatment for alcohol use	4.1	5.4	0.0	5.0	10.9	1.0	2.2	8.2
Past-year treatment for illicit drug use	4.1	5.4	0.0	5.0	13.4	1.0	2.2	10.1
Past-year treatment for drug or alcohol use	4.6	6.5	0.0	5.8	17.6	8.2	19.1	13.3
Past year treatment for drug and alcohol use	4.1	5.4	0.0	5.0	10.9	1.0	2.2	8.2
Age 14–15								
Lifetime treatment for drug or alcohol use	9.7	10.1	12.3	9.8	16.4	19.6	34.2	15.8
Past-year treatment for alcohol use	6.0	6.8	10.9	6.1	8.1	12.5	28.7	7.8
Past-year treatment for illicit drug use	6.9	3.3	11.3	4.3	9.6	10.0	25.1	7.7
Past-year treatment for drug or alcohol use	9.2	9.6	12.3	9.2	13.2	14.4	34.2	11.0
Past-year treatment for drug and alcohol use	6.0	2.0	10.9	3.1	6.6	9.4	25.1	5.7
Age 16–17								
Lifetime treatment for drug or alcohol use	12.6	11.3	8.4	12.3	22.4	20.4	32.3	20.0
Past-year treatment for alcohol use	7.2	5.0	8.4	5.6	9.1	13.6	17.2	10.3
Past-year treatment for illicit drug use	7.5	6.8	8.4	6.9	15.3	11.5	21.6	12.5
Past-year treatment for drug or alcohol use	9.0	8.4	8.4	8.7	17.5	17.6	26.0	16.3
Past year treatment for drug and alcohol use	6.7	4.3	8.4	4.9	8.9	10.3	17.2	8.5
Age 18–20								
Lifetime treatment for drug or alcohol use	11.4	12.8	9.9	12.7	21.3	16.8	30.0	17.4
Past year treatment for alcohol use	4.6	5.9	2.7	5.8	9.8	10.4	18.7	8.7
Past-year treatment for illicit drug use	3.3	3.4	4.2	3.4	11.4	6.0	16.2	7.6
Past-year treatment for drug or alcohol use	6.9	7.5	4.8	7.6	14.7	11.0	19.8	11.9

(*Continued*)

TABLE 57.10 Percentage Receiving Treatment by Substance for All AUD and SUD Individuals Ages 12–34, NSDUH 2009—cont'd

	Past-year abuse				Past-year dependence			
Substance treatment	**Illicit drugs**	**Alcohol**	**Both illicit drugs and alcohol**	**Illicit drugs or alcohol**	**Illicit drugs**	**Alcohol**	**Both illicit drugs and alcohol**	**Illicit drugs or alcohol**
Past-year treatment for drug and alcohol use	1.4	2.5	2.2	2.2	7.3	5.4	15.1	5.0
Last/current treatment for illicit drugs	2.3	2.4	2.8	2.4	10.7	5.5	16.2	6.9
Age 21–25								
Lifetime treatment for drug or alcohol use	18.5	12.9	18.2	13.6	31.3	24.7	36.7	26.2
Past-year treatment for alcohol use	8.7	4.6	8.2	5.1	11.3	12.2	20.5	10.6
Past-year treatment for illicit drug use	9.4	3.8	9.9	4.4	16.3	7.9	22.0	9.9
Past-year treatment for drug or alcohol use	10.5	5.3	10.9	5.9	17.9	14.3	25.3	14.5
Past-year treatment for drug and alcohol use	7.7	3.2	7.1	3.8	10.3	7.1	18.5	7.0
Age 26–34								
Lifetime treatment for drug or alcohol use	24.6	19.2	15.2	20.4	48.3	27.0	47.2	32.4
Past-year treatment for alcohol use	6.6	4.7	1.5	5.3	12.5	13.8	26.9	12.0
Past-year treatment for illicit drug use	3.3	1.4	1.5	1.7	15.4	6.7	25.1	7.9
Past-year treatment for drug or alcohol use	6.9	5.1	1.5	5.6	20.5	17.1	32.6	16.7
Past-year treatment for drug and alcohol use	3.0	1.4	1.5	1.6	8.0	4.9	19.5	4.5

Data from the public use data set of the National Survey on Drug Use and Health conducted by the Substance Abuse and Mental Health Services Administration.

perceived availability of cannabis. However, the students in Spain and the Czech Republic report the highest 30-day use of marijuana or hashish. Reporting the highest levels of lifetime illicit drug use other than marijuana and hashish are the students from the United States (18%) and the Isle of Man (16%) followed by France, Austria, and Latvia (all at 11%).

A CONTINUAL CHALLENGE – RECENT AND EMERGING ISSUES

While substance use and abuse among adolescents and young adults is not a new phenomenon, what substances are used and how they are used often reflects what is popular among youth at a particular moment in time. Examples include drinking games, the custom of drinking 21 shots to celebrate one's 21st birthday, and the unprescribed use of medication for attention deficit hyperactivity disorder to enhance academic performance. Fishing/pharming parties are a relatively new and particularly dangerous practice in which partiers contribute whatever is available from the medicine cabinet at their home to a common "punch" bowl and everyone downs a fistful of pills. Drug and drug/alcohol interactions may be a particularly lethal combination that results.

Recently and of increasing concern, the prevalence of prescription drug abuse in the adolescent and young adult population has been increasing. This is reflected, for example, in the fact that the number of emergency department visits related to nonmedical use of opioid pain relievers increased by nearly 80% for 17-year-olds and more than doubled for 18–20 and 21–24-year-olds from 2004 to 2008 (NSDUH, data not shown).

WHAT DOES IT ALL MEAN? SUMMARY AND FINAL THOUGHTS

We know from this overview that many adolescents in the United States use some substances at some point during their development, the majority continue this use into young adulthood, and many adolescents and young adults not only use substances but also do so frequently and/or at high levels.

While we can list many of the potential consequences of substance use, we know far less about the actual numbers of specific consequences that do occur. Especially difficult to measure are the effects on long-term outcomes. For example, it is virtually impossible to enumerate how many downstream problems are set in motion by substance use or to determine the number of individuals for whom substance use has either temporarily or permanently interfered with their developmental trajectories.

What we can take away from this epidemiologic overview is some sense of the complexity of substance use among young people and the extent to which it is intertwined with development in general, and with psychiatric and other problems, and risk behaviors, specifically. Clearly, substance use is a significant problem among adolescents and young adults, affecting both boys and girls, and young people of every ethnicity, no matter where they live. One can also readily appreciate that far too many youth and young adults have experience with substances, sometimes more than one, and that they are suffering the consequences of that use including dependence. We as a society need to change attitudes so that substance use among young people becomes far less acceptable to young people themselves, as well as to the rest of us. This will be difficult as long as substance use is seen as a statement of maturity and/or identity, as central to socializing, or as a way to escape difficulties. The ultimate goal is to help young people navigate the transition from childhood to adulthood without the use of substances which can hijack promising young lives.

List of Abbreviations

CDC Centers for Disease Control and Prevention
ESPAD European School Survey Project on Alcohol and Other Drugs
HIV human immunodeficiency virus
LSD lysergic acid diethylamide
MTF monitoring the future
NSDUH National Survey on Drug Use and Health
PCP Phencyclidine
SUDs substance use disorders
YRBS Youth Risk Behavior Survey
YRBSS Youth Risk Behavior Surveillance System

Further Reading

U.S. Department of Health and Human Services, 2007. The Surgeon General's Call to Action to Prevent and Reduce Underage Drinking. U.S. Department of Health and Human Services, Office of the Surgeon General.

Faden, V.B., Goldman, M.S. (Eds.), April 2008. Underage drinking: understanding and reducing risk in the context of human development. Pediatrics 121 (Suppl. 4).

A developmental perspective on underage alcohol use. Alcohol Alert 78. National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, US Department of Health and Human Services, July 2009.

Arnett, J.J., 2000. Emerging adulthood: a theory of development from the late teens through the twenties. American Psychologist 55 (5), 469–480. PMID: 10842426.

Grant, B.F., Dawson, D.A., 1998. Age of onset of drug use and its association with DSM-IV drug abuse and dependence: results from the National Longitudinal Alcohol Epidemiologic Survey. Journal of Substance Abuse 10 (2), 163–173.

Hingson, R., Zha, W., 2009. Magnitude of and trends in alcohol-related mortality and morbidity among U.S. college students ages 18–24, 1998–2005. Journal of Studies on Alcohol and Drugs Supplement 16, 12–20.

Hingson, R.W., Heeren, T., Winter, M.R., 2006. Age at drinking onset and alcohol dependence: age at onset, duration, and severity. Archives of Pediatric and Adolescent Medicine 160, 739–746.

Masten, A.S., Faden, V.B., Zucker, R.A., Spear, L.P., April 2008. Underage drinking a developmental framework. Pediatrics 121 (Suppl. 4), S235–S251.

Relevant Websites

http://www.cdc.gov/injury/wisqars/index.html – Centers for Disease Control and Prevention, Injury Prevention and Control.
http://www.cdc.gov/Features/RiskBehavior – Children & Health, Immunization, Data & Statistics, Wildfire safety, Environmental Health.
http://www.espad.org/espad-reports – European School Survey Project on Alcohol and Other Drugs.
http://monitoringthefuture.org – Monitoring the Future survey.
http://www.niaaa.nih.gov – National Institute on Alcohol Abuse and Alcoholism.
http://www.nida.nih.gov – National Institute on Drug Abuse.
https://nsduhweb.rti.org – National Survey on Drug Use and Health.
http://www/samhsa.gov – Substance Abuse and Mental Health Administration.
http://www.cdc.gov/HealthyYouth/yrbs/index.htm – Youth Risk Behavior Surveillance System and Survey.

CHAPTER

58

Alcohol and Drug Use in Lesbian, Gay, Bisexual, and Transgender (LGBT) Youth and Young Adults

*Kimberly F. Balsam**, *Yamile Molina*[$], *Keren Lehavot*****

[*]Palo Alto University, Palo Alto, CA, USA [$]Fred Hutchinson Cancer Research Center, Seattle, WA, USA
[**]VA Puget Sound Healthcare System, Seattle, WA, USA

OUTLINE

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00058-9

INTRODUCTION

Lesbian, gay, bisexual, and transgender (LGBT) populations are at elevated risk for substance use and abuse compared to their heterosexual counterparts. This chapter focuses on alcohol and drug use, abuse, and dependency among LGBT populations with a particular emphasis on the developmental period of youth and young adults. We begin by defining sexual orientation and gender identity, discuss the unique life challenges faced by LGBT individuals, and describe how these challenges are linked to substance use, abuse, and dependency for LGBT populations in general. We proceed to discuss some of the challenges that are particular to LGBT youth and young adults, provide details about the use and abuse of specific substances in LGBT youth populations, and discuss potential explanations for elevated use in this population. We conclude the chapter with an overview of substance abuse treatment issues specific to this high-risk population.

OVERVIEW OF LGBT POPULATIONS

Definitions and Dimensions of Sexual Orientation

Sexual orientation is a complex, multidimensional construct for which definition and measurement can vary across studies and settings. However, there is a growing consensus that sexual orientation includes cognitive, affective, and behavioral dimensions. *Sexual identity,* the cognitive dimension of sexual orientation, is the term typically used to refer to an individual's own self-identification as lesbian, gay, or bisexual (LGB). In contrast, behavioral definitions of sexual orientation focus on the sex of an individual's sexual partners. Research using behavioral definitions sometimes refers to participants as MSM (men who have sex with men) or WSW (women who have sex with women). Sexual attraction, the affective component of sexual orientation, can also be assessed. Individuals may be emotionally and sexually attracted to men, women, or both. Although these three dimensions of sexual orientation are related, there is not complete overlap among them. Indeed, research that assesses all three dimensions of sexual orientation reveals a surprising lack of congruence, particularly among youth and women. For example, HIV and AIDS researchers must contend with the fact that some men identify outwardly as heterosexual yet may engage in sex with other men covertly. Similarly, some women may experience attractions to both women and men yet identify as lesbian because of the stigma associated with a bisexual identity in lesbian communities.

Gender Identity

Gender identity is a term used to refer to an individual's internal identification with being male or female. In contrast to sexual identity, gender identity is focused on one's view of oneself, rather than one's attractions to others. Although it is often conflated with sexual identity in the literature, gender identity is a distinct construct. *Transgender* is an umbrella term that encompasses a range of variations in gender identity, including transsexuals (individuals who identify as or desire to live and be accepted as a member of the sex different from the one they were assigned at birth), transvestites (individuals who cross-dress), and individuals whose gender identity lies outside of the traditional binary categorization of gender (e.g. genderqueer, two-spirit). Transgender individuals may include male-to-female (MTF) and female-to-male (FTM) individuals. Transgender individuals may identify their sexual identity as LGB, or heterosexual.

Terms for this Chapter

Although gender and sexual identity are distinct constructs, there is much overlap in the types of discrimination, prejudice, and stigma faced by lesbian, gay, bisexual and transgender people. As a result, the term "LGBT" is often used to refer to the community of sexual and gender minority people. To facilitate readability, we will use this term throughout the chapter when discussing the field more generally. However, research studies may use varying definitions of sexual minorities such as men who have sex with men (MSM) or may include only LGB (but not transgender) participants. In order to report the findings of these studies accurately, we use these terms in describing the results of specific studies. Additionally, throughout the chapter we use the following terms for racial/ethnic identities: White, Black, Latino, Native American/American Indian, Asian American, and Asian American/Pacific Islander (wherever applicable).

The Coming Out Process: Development and Milestones

One of the most central experiences for LGBT individuals is the process of *coming out,* a term used to describe the process of developing and disclosing an identity as an LGBT person. Individuals often progress through a series of milestones during this process, including awareness of same-sex attractions, self-identification as LGB, disclosure of sexual identity to others, engaging in same-sex sexual exploration, and entering into same-sex relationships. The coming out process for bisexuals may be more complex, as individuals develop

awareness of both same- and opposite-sex attractions. For transgender individuals, coming out involves awareness of gender identification, disclosure of this identification to others, and for some, changing one's outward gender expression and/or physical body. An important step in the coming out process is seeking and finding support from other LGBT individuals and communities. The timing and sequence of coming out milestones can vary greatly according to the individual's social and cultural context as well as their own individual characteristics and personal history. It is also important to note that disclosure of sexual orientation is an ongoing process over the lifespan, with each new interpersonal encounter presenting challenges of when and how to come out. Individual's degree of "outness" may also vary according to context; for example, a gay man may decide not to disclose his sexual orientation at work, but may be very open with his friends and family about his identity and relationships.

OPPRESSION OF LGBT PEOPLE

The term "homophobia," which refers to the irrational fear, hatred, and intolerance of homosexuality, has historically been used when describing the oppression faced by LGBT people in society. However, the recognition that prejudice is a social-level phenomenon, rather than an individual pathology, has led to the more common use of the term "heterosexism," which refers to the ideological denial, denigration, and stigmatization of sexual minorities. Despite much progress over the past few decades, heterosexist laws, policies, attitudes, and behaviors are still widespread in the United States. For example, federal law does not protect employees from discrimination based on sexual or gender identity, and several states do not include these identities in hate crime laws. Same-sex couples are still unable to marry or form legal unions in most parts of the United States; over two dozen states have recently changed their constitutions to restrict marriage to one man and one woman. Such legislation legalizes unequal treatment of LGBT individuals and families and exacerbates their stigmatized status.

The impact of the heterosexist social context on the individual LGBT person has been conceptualized as "minority stress," a term used to describe the additional stress exposure experienced by members of minority groups. Minority stress for LGBT people includes actual experiences of victimization and discrimination, expectations of rejection, concealment of identity, and internalized negative beliefs about being LGBT. LGBT people may also encounter minority stress due to multiple oppressed identities; for example, ethnic minority LGBTs may encounter negative attitudes regarding homosexuality in their ethnic communities, while encountering racism and rejection from other LGBT people.

A growing body of literature has documented that minority stressors are indeed common for LGBT people. A recent national probability sample of LGB adults indicated that 20% reported having experienced a person or property crime based on their sexual orientation, about half had experienced verbal harassment, and more than 1 in 10 reported having experienced employment or housing discrimination. Numerous studies have also documented that LGBT people are at elevated risk for experiences of physical and sexual violence over the lifespan. These external experiences, along with more internal minority stress processes such as internalized heterosexism (internalized negative attitudes about one's own sexual orientation), have been linked empirically to psychological distress, depression and anxiety disorders, smoking, and substance use. Indeed, minority stress has been posited to be the primary explanatory factor driving health disparities between LGBT and heterosexual populations.

SUBSTANCE USE IN LGBT POPULATIONS

Early research on LGBTs and substance use focused on lesbians and gay men; this research consistently found higher rates of substance use, abuse, and dependency relative to heterosexual adults. Because of the difficulty of finding a hidden and stigmatized population for research, these studies relied on gay bars as a venue for research recruitment, thereby introducing a confound into research findings. For example, some early studies suggested that over one-third of lesbians and gay men abused alcohol. Subsequent studies using more representative sampling methods have tempered these initially disturbing statistics. Still, population-based and other large studies of LGBTs suggest that this population does indeed use alcohol and drugs more than the general population, although these differences are greater for women than for men. It is important to note that these differences are relative and that the majority of LGBT people do not experience substance abuse or dependence.

Sexual Orientation Differences by Substance

Literature on substance use among LGBT populations has most often focused on alcohol; however, the extent to which LGBT populations differ from heterosexual populations may vary for different substances. Below we describe what is known about specific substances

and their use among LGBT populations. The majority of studies focus exclusively on LGBs, with very little empirical research on transgender populations. Additionally, samples for these studies are typically predominantly White; thus, we know relatively little about risk for substance use among ethnic minority LGBT people.

Alcohol

Alcohol may be the most commonly used and abused substance among LGBT people. Differences between LGBT and heterosexual populations are most dramatic among women. Lesbian and bisexual women use alcohol more frequently, consume greater quantities, and report higher rates of alcohol-related problems than heterosexual women. One study found that 15% of lesbians did *not* consume alcohol compared with 35% of the heterosexual women in their sample, and that lesbians were almost three times more likely to report alcohol use problems than their heterosexual peers (23% versus 8%). According to the National Lesbian Health Care Survey, 25% of a sample of 1925 women reported drinking alcohol several times a week and 6% self-reported daily drinking. Whereas hazardous drinking typically peaks in young adulthood for heterosexual women, lesbian and bisexual women continue to drink heavily into their 30s and 40s. Although early research focused primarily on lesbians, recent studies indicate that bisexual women show the highest risk for hazardous drinking.

Findings regarding gay and bisexual men's alcohol use are mixed, with most studies showing similar levels of consumption and alcohol-related problems compared to heterosexual men. However, a few studies have found high rates of abuse and dependence among gay and bisexual men. For example, one large survey of men who have sex with men (MSM) in four large metropolitan areas found that 85% reported alcohol use within the past 6 months, 12% reported three or more alcohol-related problems, and 8% reported heavy–frequent alcohol use. In comparison, the 2005–2007 National Health Interview Survey reported 67% of male respondents were current drinkers (alcohol use within the past year); DHSS data suggest that approximately 9% of men drink heavily frequently. Similarly, analyses from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC), gay men had significantly higher odds than heterosexual men of alcohol dependence within the past year.

Marijuana

Marijuana is the most commonly used and abused illicit drug in the United States, and may be even more commonly used among LGBT populations. Studies have found rates of marijuana use among lesbian and bisexual women between 14% and 38% for past year and 58% and 76% for lifetime. Comparisons with heterosexual women have revealed a greater likelihood of recent and lifetime use. Gay and bisexual men also have a greater likelihood of recent use compared to their heterosexual counterparts; however, their rates appear similar to heterosexual men for lifetime use (65.6% versus 52.3%). Lesbian, gay, and bisexual men and women in the 1996 National Household Survey on Drug Abuse (NHSDA) were more likely to meet criteria for marijuana dependence syndrome than heterosexual men and women. Similar to findings with alcohol, recent research suggests that bisexual women may be the highest risk group for marijuana use and abuse.

Methamphetamines

Research on methamphetamine use and dependency has generally focused on gay men and men who have sex with men (MSM). The past decade has seen a major surge in methamphetamine use in this population of men, with estimates of use in the United States ranging from 25% to 29%, a rate approximately 10 times higher than in the general population of men. Methamphetamine use is much less common among lesbian and bisexual women, although surveys of lifetime use indicate that lesbian and bisexual women are more likely to have tried this substance than their heterosexual counterparts.

Other Substances

LGBT individuals are also more likely to use and abuse other illicit drugs. According to 1996 NHSDA, lesbian and bisexual women reported higher lifetime rates of cocaine and hallucinogen use and abuse compared to heterosexual women. Similarly, gay and bisexual men reported elevated levels of sedative, tranquilizer, cocaine, GHB, and illicit prescription pain medication use than other men. However, research on other illicit substances is a new area of study, and thus is hampered by some of the sampling problems encountered by early research on alcohol use (e.g. recruitment of men at gay bars for research).

SUBSTANCE USE DIFFERENCES ACROSS RACE/ETHNICITY IN LGBT POPULATIONS

Though much of research on substance use in LGBT populations has utilized predominantly White samples, there has been a surge of research dedicated to individuals who are both racial/ethnic and sexual minorities. The majority of studies have addressed substance use in Black and Latino gay and bisexual men. Studies have indicated that the substances used by Black gay and bisexual men are alcohol, methamphetamine, and

crack cocaine and that Latino gay and bisexual men are more likely to use cocaine, methamphetamine, and ecstasy more often than other substances. With regard to racial/ethnic comparisons, two studies have found elevated risk among Black gay and bisexual men relative to their Latino and White counterparts, with one study focusing on illicit drug use and the other on quantity of alcohol consumption. Latino gay and bisexual men have been reported to use cocaine more often than White gay and bisexual men within the past 4 months. Other studies have indicated a higher recent (4–6 months) and lifetime rates of marijuana and crystal methamphetamine by White gay and bisexual men than Black and Latino gay and bisexual men. There are, however, a growing proportion of Black gay and bisexual men who have used methamphetamine at least once in their lifetime, and likelihood of lifetime and current (past 30 days) use has been associated with racism within predominantly White gay communities.

Substantially less is known regarding differences in substance use by race/ethnicity in lesbian and bisexual women. The majority of these studies have focused on alcohol consumption. Limited research has indicated that Black and Latina lesbians' drinking patterns are more similar to non-Latina White lesbians than to Black and Latina heterosexual women. One study assessing differences in alcohol consumption found no significant differences between Latino/Latina and non-Latino White women. There were a few differences between Black and White lesbians: Black lesbians (15.4%) were more likely to be heavy drinkers than non-Latino White women (3.8%) and were more likely to have experienced adverse consequences from drinking (e.g. driving while intoxicated, fights with family or others) within the past year.

SUBSTANCE USE IN TRANSGENDER POPULATIONS

Although public health surveillance surveys have gradually begun to include questions about sexual identity and/or same-sex sexual behavior, to date there are very few surveys that include questions pertaining to gender identity, with most surveys assessing gender as only "male" or "female." The few studies that have specifically examined transgender individuals include samples of MTF sex workers and primarily focus on HIV risk and related behaviors. Not surprisingly, rates of alcohol and illicit drug use in this population are high. However, we know much less about substance use among MTF individuals who are not sex workers, FTM individuals, and other transgender individuals (e.g. those who do not identify as either male or female).

A meta-analysis suggested sizeable use of substance use for MTFs with regard to alcohol (44%), marijuana (20%), and other illicit drugs (26.7%), but a low rate of injecting drugs such as heroin or crack. Findings are mixed with regard to FTMs: whereas two studies have indicated no injection drug use among FTMs in their small research samples, three other studies have found that 4–21% of participants reported injecting substances at least once in their lifetimes. A sample of 48 transgender youth and adults revealed that participants believed stimulants and psychedelic drugs to be the most frequently used substances in transgender communities; other research has also suggested a high rate of stimulant use by transgender populations.

CORRELATES OF SUBSTANCE USE IN LGBT POPULATIONS

Given the consistent findings of relatively high levels of substance use, abuse, and dependency among LGBT populations, recent scholarship has begun to focus on identifying potential explanations for this disparity. Many risk factors may be similar to those in the general population (e.g. familial substance abuse); higher rates of these risk factors among LGBTs may account for some of their disparities in substance use. Other risk factors, described below, may be unique to LGBT populations.

Minority Stress

Stress exposure is an important predictor of substance use in the general population. In addition to general stressors, LGBT people are subject to unique minority stressors that may lead to substance use. Perceived discrimination and harassment based on sexual orientation have been linked empirically to substance use and substance use disorders (SUDs) among LGB adults. Further, there is some evidence that internalized heterosexism may be related to hazardous drinking among lesbian and bisexual women. Substance use may be an attempt to cope with psychological distress caused by both external and internal minority stressors. Although more research is needed to better understand these causal pathways, one recent longitudinal study found that the relationship between discrimination and problem drinking was indeed mediated by both negative effect and coping motives in stigmatized populations, including women, racial/ethnic minorities, and LGB individuals.

Victimization

As previously noted, LGBTs consistently report alarmingly high rates of physical and sexual abuse and

assault over the lifespan. Men and women with victimization histories in the general population are at elevated risk for substance use and abuse, and this appears to be true for LGBT people as well. A recent analysis of data from NESARC indicates that childhood victimization may be the strongest predictor of SUDs in adulthood for LGB people. Additionally, childhood victimization is a risk factor for victimization in adulthood. This phenomenon is referred to in the literature as *revictimization* and is often used to refer to experiences of both childhood sexual abuse and adulthood sexual assault. Evidence suggests that individuals who experience revictimization are at even greater risk for substance use than those with only one type of victimization, and this is equally true for LGB individuals. Causal pathways between substance use and victimization may be bidirectional. For example, substance use may be an attempt to numb oneself to the psychological aftermath of victimization; at the same time, substance use may place individuals in situations and contexts in which victimization is more likely to occur.

Mental Health Problems

Mental health disorders such as depression and anxiety are highly comorbid with SUDs in the general population, particularly among women. It is not surprising, then, that these comorobidities also occur in LGB populations. Numerous studies have documented small but significant elevations in rates of mental health disorders and levels of psychological distress among LGB people. Mental health disorders may manifest differently among LGB people; for example, there is evidence that LGBs are even more likely to have more than one mental health or SUD in their lifetime, and are also more likely to have mental health disorders that persist over time.

LGBT-Specific Factors

Factors such as mental health problems and stressful events are hypothesized to lead individuals to "self-medicate" with substance use. However, this hypothesis does not account for substance use and abuse among individuals who are not experiencing psychological distress. Substance use, particularly alcohol use, is seen as normative in many segments of US society, including LGBT communities. Historically, gay and lesbian bars were the only public place where LGBT people could go to socialize, meet potential friends and dating partners, and freely express their identities without fear of retribution. Even with decreasing societal stigma and the resulting increase in opportunities for meeting other LGBT people, bars remain a primary socializing outlet for many LGBT people. This may be particularly true for lesbian and bisexual women, who have been documented in research to spend more time in bars than heterosexual women. Indeed, one study found that reliance on bars as a primary social venue was a predictor of substance abuse among lesbians. More recently, circuit parties (large, multi-day dance parties) have become commonplace as a socializing venue for gay and bisexual men, particularly in large cities. These events are characterized by use of multiple substances including MDMA, ketamine, methamphetamines, and cocaine. In addition to higher overall rates of illicit substance use, attendees of these parties also commonly ingest multiple substances simultaneously, including combinations of methamphetamines, ketamine, and GHB as well as combinations of club drugs with viagra. A recent study conducted in San Francisco reported a median number of drugs to be four per individual. Importantly, variation exists in individuals, wherein substance may be considerably low during non-circuit party weekends and considerably higher during circuit party weekends.

The prevalence of heavy substance use contexts in LGBT communities not only provides further opportunity to use alcohol and other drugs, but may also influence individual perceptions of social norms. LGBT people may view alcohol and other drug use as commonplace among others who share their minority status. Perceived social norms have been documented to be a strong predictor of substance use in the general population. Although there is little research on social norms in LGBT communities, preliminary evidence suggests that these norms may indeed play a role.

Another difference between LGBT and heterosexual populations that may affect substance use is life roles and transitions. The majority of heterosexual men and women marry and have children, and these life transitions are typically associated with reduction in use of substances. Although many LGBT people form and maintain long-term, intimate partnerships, legislation prohibiting same-sex marriage as well as the stigma associated with same-sex relationships serve as barriers to doing so for some LGBT individuals. Further, although the past two decades have witnessed a "gayby boom" in which more out LGBT people are choosing to become parents, the majority of LGBT people are still childless. This may be a particularly relevant factor for lesbian and bisexual women, given the sharp decreases in alcohol use typically seen among heterosexual women as they begin to have children.

In general, substance use is seen as more normative for men than for women. Hence, nontraditional gender roles and gender expression among lesbian and bisexual women may also be associated with substance use. Indeed, in one of the few studies to examine this

hypothesis, butch-identified lesbians (i.e. those adhering to a more masculine gender role and identity) were more likely to use and abuse substances than femme-identified women (i.e. those adhering to a more feminine gender role and identity). Additionally, in a recent study of gay men, those who conformed to traditional masculine gender roles were more likely to abuse alcohol and use tobacco and illicit drugs than gay men with lower masculinity. Thus, it may be that identification with traditional constructions of masculinity may be a risk factor for both men and women.

HIV

A recent study indicated that HIV-positive individuals appear to use alcohol and drugs more frequently than HIV-negative individuals in general, though other research has indicated that HIV-infected populations tend to reduce the frequency of injection drug use post-diagnosis. HIV-positive gay and bisexual men appear to be more susceptible to both risky sexual behaviors and substance use behaviors and disorders than HIV-negative gay and bisexual men. Specifically, HIV-positive men have been found to have elevated rates of heroin, crack, cocaine, and marijuana use and abuse and to be more likely to use injection drugs. Substance use in HIV-positive populations has been linked to problems with HIV treatment adherence.

SUBSTANCE ABUSE TREATMENT FOR LGBT POPULATIONS

Given their relatively high risk for substance use problems over the lifespan, it is not surprising that LGBT people are overrepresented among treatment-seekers. For example, research indicates that lesbians are more likely than heterosexual women to have engaged in treatment for alcohol and drug use problems: approximately half of lesbian and bisexual women substance users have sought professional help for their drug/alcohol problems at least once. Though research has suggested that gay and bisexual men are less likely to receive treatment relative to lesbian and bisexual women, gay and bisexual men are up to three times more likely to be receiving treatment for substance use than heterosexual men. Another study revealed that approximately 16% of gay and bisexual men reported having received treatment at some point in their lives relative to 2% of heterosexual men.

Despite demonstrated interest and need, there are relatively few programs specifically tailored for LGBT clients. Although the substance abuse field has increasingly been adapting services for culturally diverse clients, these efforts typically do not include a focus on LGBT issues. A recent study of substance abuse treatment programs in the United States and Puerto Rico found that only 11.8% provided services specifically for LGBT individuals, and only 7.3% offered tailored programs to meet their unique needs. Surveys of substance abuse treatment providers and staff reveal little to no formal education regarding the needs of LGBT clients. Moreover, these surveys reveal that many providers harbor negative or ambivalent attitudes toward LGBT clients. These attitudes do not go unnoticed; indeed, a survey of LGB adults in recovery found that most had experienced nonaffirming behaviors from treatment providers. Thus, although many LGBT people seek substance abuse treatment, it is likely that the lack of availability of culturally tailored treatment and culturally sensitive providers may hamper treatment retention and effectiveness.

Greater efforts to adapt and develop specialized services that are culturally appropriate for LGBT populations is greatly needed, considering that preliminary research has found that LGBT specialized treatment programs may be more effective than traditional, general programs for this population. Further, many LGBT individuals who use substances report a preference for an LGBT-only and/or culturally adapted treatment program.

LGBT YOUTH AND YOUNG ADULTS

The remainder of this chapter will focus on the unique factors pertaining to substance use for LGBT people in adolescence and young adulthood. This developmental period is characterized by the opportunity for identity exploration across multiple domains, including romantic relationships, academic and vocational interests, and sense of self. For many LGBT people, this is the period in which they first become aware of and disclose their LGBT identity and initiate same-sex relationships. This identity exploration and formation occurs against a backdrop of the more typical stressors faced by youth.

Youth and young adulthood is also a period characterized by increased vulnerability to substance use and abuse. Indeed, this phase represents the highest risk period for lifetime diagnoses of SUDs. Youth and young adults are particularly vulnerable to negative outcomes associated with substance use. For example, alcohol use is the largest contributor to preventable deaths in this age group, associated with the three leading causes of mortality: accidents, homicide, and suicide. LGBT youth, who face additional developmental tasks and challenges during this period, are even more vulnerable to substance use and its consequences.

It is important to note the methodological challenges inherent in studying substance abuse among LGBT youth. Although public health surveys of adults have begun to add questions assessing sexual orientation, the inclusion of such questions in surveys of youth is often met with political resistance, resulting in fewer opportunities to obtain data on this population. Similar to adults, LGBT youth and young adults are a stigmatized population and thus may be difficult to recruit for research. Recruitment at this developmental phase may be further limited by the fact that many LGBT youth still reside at home and their families may not know about their sexual orientation. Additionally, many youth and young adults who will eventually identify as LGBT may still be questioning their identity and may not have disclosed to others. Accordingly, many surveys of LGBT youth and young adults focus on attractions, rather than identity or behavior, to classify this group.

Substance Use in Heterosexual and LGBT Youth and Young Adults

In most cases, LGBT youth (18 years or younger) and young adults (19–25 years old) appear to be more vulnerable to substance use, abuse, and dependency than their heterosexual peers. Research has indicated sexual orientation differences in alcohol, marijuana, methamphetamines, and other illicit drugs. In particular, bisexual and questioning youth may be particularly vulnerable to using specific substances, as discussed below.

Sexual Orientation Differences by Substance

Alcohol

Alcohol is the most frequently studied substance when comparing LGBT and heterosexual youth and young adults. Among middle and high school students living at home, those who are not exclusively heterosexual (i.e. bisexual, questioning, and mostly heterosexual) are more likely to use alcohol (e.g. 61.2% versus 51.1%), use alcohol earlier in life, and do so more frequently than their heterosexual peers. Lesbian and bisexual young women are more likely, for instance, to have consumed a whole alcoholic drink before they enter college. Another longitudinal study found that while lesbians consumed more alcohol than their heterosexual peers during high school, gay men increased their alcohol use at greater rates than heterosexual men during the initial transition to college. Among college students, LGBT individuals have reported greater use of alcohol as well as greater frequencies of binge drinking and drunkenness at the beginning of their college career. Bisexual men and women appear to be particularly likely to report binge drinking and to consume large amounts of alcohol.

Marijuana

Marijuana is the most frequently used substance among youth and young adults in the general population; recent findings among high school students suggest that it may be used even more frequently than alcohol. LGB high school boys and girls are more likely to use marijuana (73.3%) than heterosexual peers (37.2%); additionally, those who use do so more frequently. Similar differences have been found among college students, with differences between LGB and heterosexual students increasing over time. Similar to findings among adults, bisexual youth and young adults may be the highest risk group for marijuana use.

Methamphetamines

LGBT youth in the United States appear to use crystal meth and other amphetamines more often than heterosexual adolescents in the past 3 months, 6 months, and over their lifetimes. For example, a pilot study using a convenience sample found that the odds of LGBT youth having ever using crystal meth was 26% higher than for heterosexual adolescents.

Other Substances

Although there is relatively less research on other illicit drug use, there is some evidence that LGBT adolescents are more likely to inject drugs, use cocaine crack, and take hallucinogens and ketamine than their heterosexual counterparts. Use of other drugs such as heroin, peyote, or GHB has not been found to be more highly prevalent among LGBT youth. Similarly, club drug use among young adults does not seem to differ by sexual orientation.

Sub-Group Differences Within LGBT Youth and Young Adult Populations

Similar to adult LGBT populations, substance use differences between LGBT and heterosexual youth and young adults are most dramatic and common among females. Lesbian and bisexual girls and young women are more likely to report alcohol use and alcohol-related risk behaviors, and have a drug abuse diagnosis than their heterosexual counterparts. Findings regarding sexual orientation disparities in substance use for boys and young men have been mixed. While some studies have found elevated rates of alcohol use among gay and bisexual young men, others have failed to detect differences from heterosexual young men. There is relatively little empirical information about transgender youth and young adults. This developmental phase is

considered to be particularly relevant for transgender individuals, as physical changes associated with puberty, which may further enhance discrepancies between self-identity and physical appearance, may lead to intense distress and risk for substance use problems.

Another important subgroup is those who identify as bisexual, questioning, or mostly heterosexual. A number of studies have documented elevated use of substances among these youth who do not identify as exclusively lesbian or gay. These youth have additionally been noted to display a unique pattern of substance use wherein they use substances more often at the time of first initiation and are more likely to binge drink.

Among general populations, Black and Latino youth have generally been found to consume less alcohol and use drugs less frequently than their White counterparts. The few studies that have examined racial/ethnic differences in LGBT youth and young adolescents have revealed similar patterns. For instance, in one study, White LGB youths reported more frequent consumption of alcohol relative to Black, Latino or other LGB youth and young adults of color. Although there is little research examining sexual orientation differences within racial/ethnic minority groups, one study found that Asian/Pacific Islander LGB adolescents had use patterns more similar to their White LGB counterparts than to heterosexual API adolescents.

Correlates of Substance Use in LGBT Youth and Young Adults

Victimization

LGBT youth and young adults experience elevated risk for victimization across multiple contexts that likely accounts for some of their substance use and abuse. In particular, victimization and bullying in school settings has been associated with both psychological distress and substance use. One study of high school students in the United States found that individuals who experienced homophobic teasing were more likely to use alcohol and marijuana than individuals who did not experience homophobic teasing. Interestingly, students who were questioning their sexual identity and experienced high levels of homophobic teasing also reported the highest levels of substance use, suggesting that these youth may lack some of the supports that individuals with a more clear sexual minority identity experience.

Coming Out and Minority Stress

Youth and young adults who are exploring and disclosing their identities as LGBT may be subject to unique minority stressors. Unlike adults, who have more relative control over their social and work environments, most youth have little personal choice over these environments and are often subject to home, school, and community environments that do not support their LGBT identities. Indeed, research has shown that earlier timing of awareness and disclosure of LGBT identity is a predictor of substance use, perhaps reflecting the greater vulnerability to minority stressors at younger ages. Additionally, even in families that will eventually be accepting of LGBT identities, initial reactions to disclosure are often negative, and such reactions have been linked empirically to alcohol and marijuana use in several studies. Not surprisingly, internalized heterosexism has also been linked to binge and heavy drinking among gay and lesbian youth.

Family of Origin

For all youth and young adults, parental support can serve as a buffer against the effects of life challenges. This may be particularly true for LGBTs; indeed, parental support has been documented to play an important role for those who encounter sexual orientation-based victimization. For instance, according to the Growing Up Today Study, LGB students who experienced a moderate frequency of homophobic teasing and the greatest amount of parental support reported less substance use than those who received the highest frequency of homophobic teasing and had the lowest amount of parental support. Another large-scale study indicated that substance use differences between LGB and heterosexual youth may be accounted for by lower parental support among LGB youth.

Homelessness

LGBT youth are disproportionately represented among homeless youth populations, comprising an estimated 40% of adolescents living on the street. LGBT youth may run away or be asked to leave their family home due to conflict with parents over sexual orientation. Homelessness is a known risk factor for substance use, abuse, and dependency, and this may be particularly true among LGBT youth. One study found that having recently left home was associated with higher alcohol or drug use for LGB youth.

Social Norms and Peer Relationships

Peer relationships and social norms have been found to be the strongest predictors of substance use in general youth and young adult populations, due to the importance of peer influences during this developmental phase. Peers may play a particularly important role for LGBT youth, who are seeking connection with other LGBT people in the face of rejection by family, friends, and communities. Paradoxically, by seeking this support and connection, LGBT youth may expose themselves to greater risk. For example, frequency of participation in

LGBT events and perceived connection to LGBT communities have been found to be among the strongest predictors of substance use among LGB youth, both cross-sectionally and longitudinally. Similarly, a study of college students found that those attending schools with more LGB resources engaged in more drinking. Further, a retrospective study of lesbians found that exposure to heavy drinking contexts during the coming out period was an important risk factor for later drinking behaviors.

Social norms may also play an important role for LGBT youth and young adults. Research with general samples has found that perceived norms for drinking behavior are often inflated in this age group, and that it is perceived norms, rather than actual norms, that most strongly influence drinking behavior. Indeed, a few studies of college students have found perceived levels of drinking among others to play an important role in alcohol consumption and binge drinking for lesbian and bisexual women.

Prevention and Treatment with LGBT Youth and Young Adults

Clearly, many LGBT youth and young adults would benefit from substance abuse services, both preventive and treatment-oriented. However, a number of barriers exist that may limit access to these services for this high-risk population. LGBT youth and young adults may harbor concerns about heterosexism from health care and treatment providers and may avoid disclosing their LGBT identity as a measure of self-protection. One recent study of high-school LGB youth found that only 35% had disclosed their sexual identity to their primary care physician, and even fewer bisexual youth had disclosed. When asked what physicians might do to increase the likelihood of disclosure, the majority of youth indicated that simply asking about sexual orientation would be sufficient.

Although specialized services are nonexistent in most regions of the United States, some US cities now have health centers that provide services targeted to LGBT youth. Many of these services focus on the most vulnerable segments of the LGBT community, such as homeless youth. Little or no research has examined the acceptability and effectiveness of substance abuse services for LGBT youth, either tailored or nontailored. Both SAMSHA and the Centers for Disease Control (CDC) offer guidelines for assessment and treatment with LGBT youth. These suggestions include asking questions at intake to assess all three dimensions of sexual identity as well as gender identity, assessing social environments and social contexts, and emphasizing safety and support in treatment.

Importantly, research on substance use among LGB youth suggests that prevention efforts aimed at reducing use during this phase could have long-lasting effects that extend into young adulthood. Prevention efforts, along with treatment approaches, should take into account the factors that have been found to influence substance use in this population, including victimization, minority stress, disclosure, family support, and peer norms. Additionally, it is critical that providers who work with LGBT youth in general health care and social service settings continually assess for substance use and abuse and make referrals when necessary.

SUMMARY AND CONCLUSIONS

In sum, LGBT youth and young adults are a population with elevated risk for substance use, abuse, and dependency. In addition to the typical developmental tasks of this life period, these youth face additional challenges that can influence their substance use behavior and set the stage for later problems with alcohol and drugs. More research is needed on subgroups such as bisexual, transgender, and LGBT ethnic minority youth. Additionally, it is important to understand why and how some LGBT youth resist substance use during this crucial developmental period. Further, it will be very important to identify ways to increase access to preventive and treatment services and to develop and test programs to address substance use in this population.

SEE ALSO

Alcohol Use Disorders, Marijuana Use and Abuse, Methamphetamine Addiction, Gender Differences, Interpersonal Factors and Addictive Disorders, Stress and Addiction, Minority Groups and Addictions, Developmental Risk Taking and the Natural History of Alcohol and Drug Use among Youth, Epidemiology of Adolescent and Young Adult Alcohol, Tobacco, and Drug Use and Misuse in the United States, Cultural Influences on Youth Alcohol and Drug Use, Substance Use and Mental Health Issues on the College Campus

Glossary

CDC Centers for Disease Control.

Coming out the process of becoming aware of and disclosing one's identity as lesbian, gay, bisexual, or transgender.

FTM female-to-male transgender person: an individual who was born female but identifies as male.

Gender identity an individual's internal identification as a male or female.

Heterosexism ideological denial, denigration, and stigmatization of sexual minorities.

Homophobia irrational fear, hatred, and intolerance of homosexuality.
Internalized heterosexism internalized negative beliefs about one's own sexual identity as lesbian, gay, or bisexual.
LGB lesbian, gay, bisexual.
LGBT lesbian, gay, bisexual, transgender.
Minority stress the additional burden of stress experienced by members of minority groups because of their devalued or stigmatized social identity.
MTF male-to-female transgender person: an individual who was born male but identifies as female.
MSM men who engage in sexual behaviors with men.
NHSDA National Household Survey on Drug Abuse.
NESARC National Epidemiologic Survey on Alcohol and Related Conditions.
Revictimization victimization experiences in both childhood and adulthood.
Seropositive a person whose body is producing antibodies for a disease, such as HIV.
Sexual identity an individual's self-identification as lesbian, gay, bisexual, heterosexual, or other.
Sexual orientation an individual's sexual identity, attraction to, and/or sexual behavior with members of the same and/or opposite sex.
SUDs substance use disorders.
Transgender umbrella term for individuals whose gender identity does not strictly correspond with the sex they were assigned at birth.
Two-spirit a term used in some Native American/American Indian communities to refer to a gender and/or sexual identity outside of traditional male and female heterosexual identities.
WSW women who engage in sexual behaviors with women.

Further Reading

Cochran, B.N., Peavy, K.M., Robohm, J.S., 2007. Do specialized services exist for LGBT individuals seeking treatment for substance misuse? A study of available treatment programs. Substance Use and Misuse 42, 161–176.
Cochran, S.D., Ackerman, D., Mays, V.M.M., Ross, M.W., 2003. Prevalence of non-medical drug use and dependence among homosexually active men and women in the US population. Addiction 99, 989–998.
Corliss, H.L., Rosario, M., Fisher, L.B., Austin, S.B., 2008. Sexual orientation disparities in longitudinal alcohol use patterns among adolescents: findings from the growing up today study. Archives of Pediatric and Adolescent Medicine 11, 1071–1078.
Espelage, D.L., Aragon, S.R., Birkett, M., Koenig, B.W., 2008. Homophobic teasing, psychological outcomes, and sexual orientation among high school students: What influence do parents and schools have? School Psychology Review 37, 202–216.
Hatzenbuehler, M.L., Corbin, W.R., Fromme, K., 2008. Trajectories and determinants of alcohol use among LGB young adults and their heterosexual peers: results from a prospective study. Developmental Psychology 44, 81–90.
Marshal, M.P., Friedman, M.S., Stall, R., Thompson, A.L., 2009. Individual trajectories of substance use in lesbian, gay, and bisexual youth and heterosexual youth. Addiction 104 974–961.
Marshal, M.P., Friedman, M.S., Stall, R., et al., 2008. Sexual orientation and adolescent substance use: a meta-analysis and methodological review. Addiction 103, 546–556.
McCabe, S.E., Hughes, T.L., Bostwick, W.B., West, B.T., Boyd, C.J., 2009. Sexual orientation,substance use behaviors, and substance dependence in the United States. Addiction 104, 1333–1345.
Saewyc, E.M., Bauer, G.R., Skay, C.L., et al., 2004. Measuring sexual orientation in adolescent health surveys: evaluation of eight school-based surveys. Journal of Adolescent Health 35, 345–360.
Talley, A.E., Sher, K.J., Littlefield, A.K., 2010. Sexual orientation and substance use trajectories in emerging adulthood. Addiction 105, 1235–1245.

CHAPTER

59

Athletes and Substance Use

Matthew P. Martens
University of Missouri, Columbia, MO, USA

OUTLINE

Substance use among athletes receives a considerable amount of attention in many nations, including the United States. This is presumably due to a number of factors, including the cultural importance of sport, high-profile alcohol and/or drug-related incidents involving famous athletes, and ethical issues associated with the use of performance-enhancing drugs. The impact of sport on many societies and cultures is considerable. In the United States, for example, youth and adolescent sport is pervasive, athletes are considered role-models by many, collegiate and professional sport are multibillion dollar industries, and there are dozens of television channels devoted exclusively to athletics. Considering this context, it is not surprising that substance-related issues among athletes can become magnified in terms of their overall societal impact. This impact includes tragedies such as alcohol- and drug-related deaths among famous athletes and the explosion in recent years of revelations of steroid and other performance-enhancing drug use (e.g. the large number of major league baseball players linked to steroids and human growth hormone; drug-related scandals associated with the Tour de France bicycle race). This state of affairs and the publicity afforded to substance-related issues among athletes creates the impression that substance use is rampant among this group. The reality of the situation, though, is more complex, including what drugs athletes are more likely than others to use and abuse, why it is that athletes use certain substances, and how to effectively intervene with athletes who do experience substance-related problems. Therefore, the purpose of this chapter is to review the prevalence rates of substance use among athletes, explanations for such use, and what is known about effective ways to prevent or reduce substance use in this population.

PREVALENCE AND REASONS FOR SUBSTANCE USE AMONG ATHLETES

Alcohol

The most frequently studied substance among athletes is alcohol, with the largest and most comprehensive studies occurring among college athletes in the United States. There have been three nationally representative studies that have compared drinking rates between college athletes and nonathletes, all of which have shown

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00059-0

that athletes are more likely than other college students to report heavy drinking. These studies have shown that between 53% and 57% of college athletes reported a heavy drinking or "binge" episode (defined as 5+ drinks in one sitting for men and either 4+ or 5+ drinks in one sitting for women) in the preceding 2 weeks, compared to 36–43% of nonathletes. These studies also found that athletes were more likely than nonathletes to engage in frequent heavy episodic drinking (3+ occasions in the preceding 2 weeks) and one study found that, on average, athletes consumed approximately 3.5 drinks per week more than nonathletes (7.57 drinks versus 4.12 drinks). These studies also showed that athletes were more likely than nonathletes to experience numerous negative alcohol-related consequences, such as having impaired academic performance, getting into trouble with the authorities, and experiencing unplanned or unwanted sexual encounters due to their drinking. It is therefore clear that participating in college athletics is a risk factor for excessive alcohol consumption and resulting negative consequences.

There have also been studies that compared alcohol use rates between athletes and nonathletes at the youth/high school level. In general, the findings are mixed regarding differences in alcohol consumption between youth athletes and nonathletes, although more recent studies suggest that athletes in that population drink more than nonathletes. One study of approximately 45, 000 8th, 10th, and 12th grade students in the United States found that participating in an athletic team was associated with more frequent alcohol use and increased likelihood of binge drinking. Another study using a nationally representative sample of US high school students found that participating in sport as a sole school-related activity was associated with a greater acceleration in problem alcohol use over time. However, participating in both school and academic activities was associated with a slower acceleration in problematic alcohol use. Another large-scale study among Norwegian adolescents found athletes were initially less likely than nonathletes to report being intoxicated over the previous 12 months, but experienced a greater growth than nonathletes in likelihood of intoxication over time. This effect was limited only to athletes who participated in team sports, as athletes participating in endurance sports actually had a decrease in intoxication over time relative to nonathletes. Other studies have suggested that certain subgroups of youth or high school athletes are more at risk for problematic alcohol consumption than other students, such as those with a high athletic identity and those with risk factors such as aggression and other problem behaviors.

There has been very little research examining alcohol use among athletes at the professional or other "elite" level, so little is known about the drinking habits of that group and how it compares to relevant population norms. Anecdotal reports in the popular press often suggest that professional or other elite athletes are frequently socializing and drinking heavily, but at this point there is little evidence to either support or refute this contention. There is some evidence to suggest that college students and adults participating in "nonelite" sports like intramurals or club teams are at risk for heavy alcohol use.

A few studies primarily conducted among college athletes have also shown that athletes tend to report drinking less during their competitive seasons than during the off-season. In the only longitudinal study examining this issue, the researchers found that male and female athletes reported a 69% and 58% higher rate, respectively, of typical drinks per week during the off-season in comparison to the in-season. These findings suggest that athletes are particularly at risk for heavy alcohol use and its resulting negative consequences when they are not active in their competitive seasons. It is possible that during a team's competitive season coaches are more likely to have specific rules that prohibit or limit alcohol use and teammates may be more likely to formally or informally discourage heavy drinking among their peers.

Based on this information on prevalence rates of alcohol use among athletes and nonathletes, it is important to understand why it is that some groups of athletes are more at risk relative to their nonathlete counterparts for excessive alcohol consumption and related negative consequences. Athletes are as susceptible as anyone else to the large number of risk factors associated with problematic alcohol consumption, but it is also possible that there are factors specific to sport itself that make athletes particularly at risk for heavy alcohol use. There are two main ways that researchers and theorists have attempted to explain these differences. One explanation is that athletes possess internal characteristics that make them more likely than others to engage in heavy alcohol use. The second is that athletes are susceptible to unique environmental or contextual influences that result in excessive alcohol consumption.

Personality Characteristics

An internal characteristic that may explain differences between athletes and nonathletes in their use of alcohol is differences in personality traits associated with excessive drinking. Such differences may suggest a genetic or biological risk factor for heavy alcohol consumption that is more common among athletes than nonathletes. For example, a number of researchers have shown that the personality trait of sensation seeking, which is typically conceptualized as a propensity toward enjoying exciting and/or risky activities, is associated with elevated levels of alcohol consumption. Different groups of researchers

have shown that athletes are more likely than nonathletes to report engaging in various risky behaviors and to score higher on measures of sensation seeking. Athletes in many sports experience a great deal of excitement and risk as a result of their athletic participation, which could explain why one finds higher levels of sensation seeking among athletes than others. It is therefore possible that for many athletes a personality predisposition that in part draws them to athletics also makes it more likely that they will engage in other "risky" behaviors such as excessive alcohol consumption.

Excessive Stress

A second internal characteristic that might explain drinking differences between athletes and nonathletes, which would pertain mostly to athletes at the collegiate level, is that some athletes drink heavily in response to excessive stress associated with balancing the demands of being both a student and an athlete. There is no doubt that many college athletes face a unique set of pressures, such as concerns about playing time or athletic performance and balancing academic work with 20+ h per week of sport-related commitments. Although theorists have speculated for some time that excessive stress levels are associated with alcohol use among athletes, research does not seem to suggest this is in fact the case. Studies have shown that few college athletes endorse coping-related motives as a primary reason for using alcohol and one study showed no athlete–nonathlete differences on a measure of drinking to relax. Athletes who do drink for coping-related reasons, like nonathletes who also do so, may be more at risk for significant alcohol-related problems than those who drink for reasons such as enjoying a party or experiencing pleasant feelings. But, using alcohol as a means of coping with excessive stress has not been shown to be an important reason for heavy drinking among athletes.

Social Opportunities

Relatively more attention has been given to environmental or contextual factors unique to sport that might explain higher rates of alcohol use among athletes, and it is possible that these factors interact with internal characteristics like personality traits in causing elevated risk for alcohol use among athletes. One unique environmental factor that may in part explain heavy alcohol use among athletes is the existence of more social opportunities than their peers. In many settings, such as the collegiate and high school environments, athletes are afforded a great deal of popularity and notoriety. This popularity may result in more social opportunities relative to the typical nonathlete, which for some results in excessive alcohol consumption. Athletes at the college level may be more likely than others to be invited to parties, bars, clubs, etc., where other students are interested in buying/providing them drinks due to their popularity and notoriety. Additionally, athletes on many teams have a built-in social network with teammates that can lead to opportunities for heavy drinking, especially at the collegiate level where athletes clearly drink more than other students. These social characteristics, especially if one possesses internal characteristics that make it more likely he or she will drink alcohol, could promote excessive alcohol consumption among some groups of athletes.

Athletic Team-Related Factors

Another unique environmental factor that could be associated with excessive alcohol use among some athletes is the influence of the athletic team. Social norms theory suggests that the more an individual perceives that others drink alcohol or the more permissive the attitudes others have toward drinking, the more the individual him- or herself will drink. A number of research studies have supported this contention, such that perceiving that one's peers drink heavily and/or have favorable attitudes toward drinking is associated with more alcohol use among the individual in questions. Research has also shown that the more salient or important the group that one is using in the social norms comparison (e.g. other students at one's university versus college students in general), the stronger the relationship between these perceived norms and personal alcohol use. In many instances, an athlete's teammates serve as important social influences over his or her life, so if an athlete perceives that heavy drinking is normal or encouraged by his or her teammates it is more likely that he or she will also drink heavily. Teammates may also influence each others' drinking through more overt means, such as formal or informal drinking-related competitions. Some theorists have speculated that the competitiveness that athletes display with both their teammates and opponents on the practice and playing field translates to the drinking arena as well. One study found that college athletes were more likely than other students to report engaging in drinking games, which are associated with heavy overall alcohol consumption. Drinking games involve competitions where the loser is required to drink a certain amount of alcohol. For example, in a game like beer pong one has to drink if his or her opponent successfully hits the ping pong ball into the designated cup. Thus, in an effort to "win" formal or informal competitive drinking competitions with their teammates some athletes may engage in risky drinking practices such as consuming large amounts of alcohol over a short period of time.

Other Cultural Factors

More recently researchers have begun to examine the association between broad cultural factors like alcohol

industry sponsorship of sport and its potential impact on the drinking behaviors of athletes. In Europe, pubs have been associated with sporting events for several hundred years and have historically played an important role in promoting sport. Currently, it is fair to say that alcohol companies are major sponsors of professional sports throughout the world, and in some cases have even owned professional teams or leant their names to professional stadiums. For example, the St. Louis Cardinals baseball team was owned by the Busch family for almost half a century and stadiums in places like St. Louis, Denver, and Montreal are named after beer companies. It is not clear how something like broad industry sponsorship of sport might impact the drinking behaviors of athletes themselves, but one possibility is that those who participate in sport have been influenced by advertising messages about their association at a greater intensity than others in the population. It is also possible that for some athletes this sponsorship link impacts their perceptions of what behaviors are normal or encouraged as part of being an athlete. In some cases a more direct relationship might exist between alcohol sponsorship and individual drinking behaviors. For example, one large study of athletes in New Zealand found that almost half of the athletes were sponsored by the alcohol industry, almost half of those sponsored reported receiving free or reduced price alcohol-related products, and those receiving alcohol industry sponsorship reported more problematic alcohol use than those not receiving sponsorship. In sum, there are a host of factors that might explain excessive alcohol use among athletes. It is important that researchers interested in understanding the phenomenon of alcohol use among athletes consider embarking upon high-quality studies designed to address topics reviewed in this section.

Tobacco

Research has shown that athletes generally report less cigarette use than nonathletes. National studies of US college students indicate approximately 15–18% of students overall report smoking cigarettes at least once in the past 30 days, whereas national studies of college athletes have shown only 9% report past 30-day use. Further, these studies show that the majority of college athletes who report using cigarettes indicate they do so only on social occasions with friends, whereas almost 50% of college students in general who smoke cigarettes do so daily. A national study among US high school athletes reported a similar pattern of findings, with female athletes 29% less likely and male athletes 36% less likely to smoke regularly than their nonathlete peers. Studies among other groups of athletes including Greek adolescents, French elite student-athletes, and English recreational athletes have shown that athletes were less likely than nonathlete comparison groups to report cigarette use.

One type of tobacco athletes are more likely to use, at least among some sports in the United States, is smokeless or spit tobacco. Approximately 12% of US college athletes report smokeless tobacco use in the past 30 days, compared to only 4% of college students in general. Smokeless tobacco use among athletes occurs disproportionately among certain types of sports. For example, approximately 40% of US collegiate baseball players and 30% of football players report past year smokeless tobacco use, compared to only 10–12% of those participating in basketball or track and field. Among US high school athletes, a national survey found that female athletes were 83% more likely and male athletes 41% more likely to use smokeless tobacco than nonathletes.

These findings represent an interesting discrepancy in tobacco use among athletes. The fact that cigarette use is relatively low among athletes is not surprising, due in part to the fact that athletes are more likely than others to be concerned about the health-related risks of cigarette use and/or its impact on their ability to perform as athletes. In contrast, many athletes do not seem to share the same concerns regarding smokeless tobacco. One explanation for this discrepancy is the cultural link between smokeless tobacco and certain types of sports. For example, one only need to watch a major league baseball game on television to see multiple instances of players using smokeless tobacco either on the bench or when actually playing in the game. Many US college athletes report that a member of their coaching staff or other athletic department personnel uses smokeless tobacco, which may lead to increased perceptions of permissiveness among one's coaches. Finally, athletes may be more comfortable using smokeless tobacco in comparison to cigarettes because they do not perceive it to have as strong of a negative impact on their athletic performance. Fortunately, researchers have shown that spit tobacco interventions that include dental examinations and behavioral counseling are effective at reducing tobacco use among adolescent athletes.

Illicit Recreational Drugs

The largest studies of recreational drug use among athletes have been conducted by the National Collegiate Athletic Association (NCAA). Their most recent study involved almost 20, 000 college athletes at universities across the United States. Results from the study indicated that in the past year 20.3% of college athletes reported using marijuana, 2.5% reported using hallucinogens (excluding ecstasy), 2.1% reported using cocaine,

and 1.2% reported using ecstasy. These figures are either comparable or lower to prevalence rates obtained from national surveys of the general college student population. For example, one national survey conducted in the same year as the aforementioned NCAA study found past-year prevalence rates of 33.3% for marijuana and 5.7% for cocaine among the general college student population. Additionally, a national study comparing marijuana use among college athletes and nonathletes found that male athletes were less likely than nonathletes to report marijuana use in the past 30 days, while rates were similar for female athletes and nonathletes. Findings from studies on other groups of athletes have generally reported a similar pattern of results. Studies of youth athletes in both the United States and other countries have usually shown that those participating in athletics are either less likely or as likely as those not participating in athletics to report recreational drug use, and the limited research on athletes at the professional or elite level also suggests that such individuals are less likely to use recreational drugs than nonathletes. Thus, in contrast to the findings that some groups of athletes use more alcohol than their nonathlete counterparts, research suggests that the opposite pattern occurs for recreational illicit drugs.

Even if athletes are at less or comparable risk than others for utilizing recreational drugs like marijuana, cocaine, or hallucinogens, it is still important to understand the factors associated with their use (or lack of use) specifically among athletes. In addition to the negative consequences associated with illicit drug use that can be experienced by anyone abusing a substance, some athletes who choose to use such substances experience a unique set of negative consequences associated with their participation in sport. For example, college athletes who are on scholarship risk suspension or expulsion from their sport as a result of a positive drug test, which could impact their academic futures. Athletes who have the talent to participate at a higher level, such as high school athletes wishing to participate in college athletics or college athletes wishing to compete at the professional level, may find their opportunities limited if they are known to have used illegal substances. Finally, professional athletes who are caught using illicit drugs face suspensions and other punishments that can impact their future employment and earnings potential in the form of less lucrative contracts and loss of endorsement opportunities.

One clear deterrent to recreational drug use among some athletes is random drug testing. Drug testing is commonplace among professional, college, and other elite athletes, and is instituted in some cases at the high school level. The frequency of drug testing an athlete is subjected to will vary across sport and competitive level, and even within competitive level when considering college athletes. Some colleges and universities (approximately 40%) do not engage in any internal drug testing, whereas others (generally those with more resources) will consistently randomly test their athletes. Testing is also conducted by governing organizations like the NCAA, often during championship events. Unfortunately, the effect of drug testing on athletes' substance use has not been extensively examined. One study conducted among high school athletes found some positive effects for a random drug testing program, but findings were not consistent enough to make strong conclusions about the program's effectiveness. Other studies have shown athletes are generally in favor of drug testing programs. Whether these programs are effective at preventing drug use among athletes at different competitive levels, though, awaits further study.

A second factor that may deter athletes from engaging in recreational drug use is health-related concerns. In the aforementioned survey of NCAA college athletes, the most commonly cited reason for not using recreational drugs was concerns about negative health-related consequences. This same concern was the most commonly cited reason for not using alcohol as well, but the percentage of athletes citing it as a deterrent from alcohol use as opposed to other drug use was considerably smaller. Most athletes are health-conscious, and it is possible that drugs other than alcohol are perceived by athletes as having a greater negative impact on one's health. Interestingly, concerns about a substance's negative impact on athletic performance were not frequently cited as the main reason for avoiding its use. It is therefore possible that a primary deterrent to drug use among athletes is a perception that using such substances will have considerable negative health-related outcomes.

Performance-Enhancing Drugs

Unlike alcohol and illicit recreational drug use, performance-enhancing drug use among athletes involves ethical considerations regarding competitive fairness. Competitive sport is built on principles of rules and fair play, so behavior on the part of an athlete that is deemed to create an unfair competitive advantage is often looked down upon by those involved in athletics. Even though such principles are sometimes applied unfairly and hypocritically (e.g. stealing signs in baseball is considered "part of the game" while steroid use is generally condemned), it is fair to argue that without such principles sport may cease to exist. Understanding performance enhancement drug use among athletes is therefore of interest not just because of potential health risks to the athletes who take such substances, but because of concerns about the impact of such use on the integrity of sport itself.

The consensus seems to be that performance-enhancing drug use, particularly steroids and human growth hormone, has increased over the past quarter century, and public perception may be that a large percentage of athletes use performance-enhancing drugs. Research, however, suggests that a relatively small percentage of athletes in fact use performance-enhancing drugs. The most recent NCAA study found that only 1.2% of college athletes had used steroids in the past year. A small percentage of students also reported using amphetamines (4.1%) and ephedrine (2.5%) in the previous year, but only 9.7% of those who used amphetamines and 27.3% of those who used ephedrine stated that the main reason they did so was to improve the athletic performance. There were differences across sports in terms of performance-enhancing drug use, but rates were still low among sports where one might think its use is relatively prevalent. For example, prevalence rates among baseball and football players were 2.3%, compared to 0.8% for track and field athletes and 0.3% for tennis players. Some smaller studies conducted in single universities have found higher prevalence rates of performance-enhancing drug use among college athletes than nonathletes, but overall it appears that a relatively small percentage of college athletes report using such drugs.

Studies among high school students have reported similarly low rates of steroid use among athletes, and one national study of over 16, 000 high school students found no differences between athletes and nonathletes on steroid use. No large studies have examined performance enhancement drug use rates among professional or other elite athletes. Such use is clearly occurring among some professional athletes, as evidenced by suspensions that occur each year in major professional sports and from international sporting governing bodies, and various estimates of use have been put forth in the popular press. But the actual rates of performance-enhancing drug use among professional and other elite athletes await further study.

Explaining the reasons for performance-enhancing drug use among athletes is presumably more straightforward than for other substances, in that the primary reason is to improve one's athletic performance. For example, the most commonly cited reason for steroid use among NCAA athletes was to enhance athletic performance, while assisting in recovery from injury was the second most common reason. It is possible that a number of sport-related variables serve as risk factors for steroid use among athletes, including permissive norms among teammates and coaches, a "win at all cost" mentality that is promoted within a team or organization, and believing that one is at risk for losing a spot on a team or in a starting lineup. Unfortunately, few research studies have attempted to address the role such factors might play in understanding steroid use among athletes.

TREATING AND PREVENTING SUBSTANCE USE AMONG ATHLETES

When thinking about treating and/or preventing substance use specifically among athletes, one must first consider the larger knowledge base regarding effective interventions among the general population. Substance abuse interventions can range from universal prevention strategies that are designed to be applied to an entire population in an effort to delay or inhibit the initiation of substance use to intensive treatments designed for those experiencing considerable substance-related problems. How interventions across this entire spectrum may be applied to athletes is described below.

One of the most commonly used universal prevention intervention models involve social norms marketing. This intervention has been frequently implemented on college campuses in the United States. The purpose of this type of intervention is to correct misperceptions regarding the actual rates of drinking behavior among a target population. Research has consistently shown that most college students overestimate the amount of alcohol that other students drink. Presumably, some students are in part motivated to drink in order to "fit in" with the perceived behaviors of other students. Thus, correcting this normative misperception may result in reduced drinking among those who had overestimated the amount of alcohol consumed by the typical college student. The effectiveness of social norms marketing campaigns has been the subject of some debate, and the research evidence is hampered by the existence of only a handful of well-designed studies that have addressed the issue. One large clinical trial did provide some support that these types of interventions, when implemented properly, can have small population-level effects.

Social norms marketing interventions can be targeted toward specific population groups, including college athletes. For example, instead of providing normative drinking rates for college students in general, the campaigns could use norms specific to college athletes or include information about misperceptions regarding the negative effects of alcohol use on athletic performance. Ideally these data would be specific to the location(s) where the intervention was being conducted, because research has shown a positive association between the salience of the normative reference group and personal substance use. For example, if a targeted social norms intervention were being developed for use at a specific high school, it would be best if the norms regarding perceived drinking rates, actual

drinking rates, and other related variables were collected from athletes at that high school.

Only a few studies have examined targeted social norms campaign among athletes, but there is some evidence that the interventions may be effective. One study examined the efficacy of a targeted social norms campaign among college athletes at a Northeastern university. The authors reported a decrease in alcohol use among student-athletes at the university over a 2-year follow-up after the campaign was implemented. They also found that students who reported more exposure to the targeted campaign also reported less alcohol use. Unfortunately, the study did not contain a control condition, thereby making conclusions about causality impossible. Nonetheless, a targeted social norms intervention may have promise as a universal prevention intervention among college athletes.

Effective interventions have also been developed for individuals who may be experiencing some problems associated with a substance but are not yet in need of more intensive treatment (i.e. indicated prevention). One well-studied intervention that has been applied in an indicated prevention framework involves providing an individual with personalized feedback regarding his or her substance use, sometimes in the context of a brief (15–50 min) one-on-one meeting with a facilitator. These interventions are rooted in Motivational Interviewing philosophy in that they are designed to increase motivation to change behavior by developing discrepancy between current behaviors and one's goals or values and overcoming ambivalence or resistance toward change. When delivered in a one-on-one session facilitators of these types of interventions use a number of techniques such as open-ended questions, reflective listening, and effective summaries to promote change talk within the session. Personalized feedback is often used to guide the content of these sessions, or in some cases is used as an intervention in and of itself. This feedback is developed based on an individual's responses to substance-related questionnaires, and often contains information such as how one's substance use compares to a relevant norm (e.g. how a college student's average drinks per week compares to a national or campus average for college students), a summary of negative consequences experienced as a result of substance use, other negative outcomes such as alcohol-related calories consumed or dollars spent on substances over a specific time period, and use or nonuse of strategies that have been shown to be associated with less substance use and fewer substance-related problems. Because this information is personalized, it is presumably more effective than general educational information in motivating people to change their behavior. A number of studies among groups such as college students, primary care patients, and emergency room patients have shown that these types of interventions are effective at reducing substance use. The majority of these studies have been in the area of alcohol use, although positive effects for substances like marijuana have also been demonstrated.

These personalized feedback-based interventions may be particularly appealing for use among athletes in part because the content of the intervention can be targeted specifically for the population. One group of researchers have shown that a standard personalized feedback-only intervention (i.e. one that was not targeted specifically toward athletes and did not involve a one-on-one meeting with a facilitator) was effective at reducing alcohol use among samples of college athletes over a 3-month period. In one of their studies, the authors reported that among athletes who had recently engaged in binge or heavy episodic drinking, those who received the personalized feedback intervention reported an average decrease of 32–46% in alcohol use (depending upon the specific alcohol use measure), whereas those in the control condition reported an average increase of 6–11%. A different group of researchers examined whether targeting personalized feedback specifically for college athletes would be more effective than more general personalized feedback that could be applied to all college students. They found that the targeted personalized feedback was more effective than the general personalized feedback and an education-only control condition at reducing alcohol use, particularly in terms of peak blood alcohol content on one occasion over the past month. A third group of researchers found that an intervention that utilized personalized normative feedback that was specific to college athletes was associated with decreased alcohol use, although the design of the study precluded causal conclusions. Finally, another group of researchers reported that an intervention that combined a brief one-on-one personalized feedback-based intervention with an intervention designed to facilitate parental communication about alcohol use was effective at reducing alcohol use among high school athletes making the transition to college. Together these studies suggest that personalized feedback-based interventions can be effective at reducing alcohol use among athletes, and that targeting the contents of interventions to be more relevant for athletes may yield additional benefits.

There are several treatments available for athletes who are experiencing significant problems with substance use and may require more intensive interventions than those used in the context of indicated prevention. The efficacy of these treatments has not been specifically examined among athletes, but there is no reason to suspect that the treatments would not generalize to the athlete population. Cognitive-behavioral and behavioral treatments that focus on factors such as developing social environments that do not reward

substance use and skills for avoiding or minimizing substance use, and contingency management treatments that provide tangible rewards for not engaging in substance use, have been shown to be effective at reducing alcohol and/or drug use. Additionally, therapies designed to facilitate the 12-step approach and marital/family treatments have been shown to reduce substance use, and pharmacological therapies like naltrexone and acamprosate have received some support. Thus, there are a number of empirically supported options that exist for the athlete who is in need of formal treatment for a substance use disorder.

When considering any type of intervention program that will be implemented or evaluated among a sample of athletes, it is important to consider several factors associated with athletic teams and organizations. First, many athletic teams, departments, and/or organizations are fairly closed systems and can be distrustful of outsiders, particularly when it comes to sensitive topics like substance use. This suggests that those interested in intervening in such a system may have to initially spend time gaining the trust of those in charge of the system (e.g. athletic directors, head coaches). Second, it is important to recognize that in most teams a coach wields a considerable amount of power and influence over his or her athletes. In many cases coaches can make arbitrary decisions regarding things like playing time, starting status, and even membership at all on a team. Although researchers have not yet examined this premise, it may be possible to work with coaches to establish clear policies and expectations regarding a team's behavior with respect to prohibition or responsible use of alcohol. Finally, considering the importance of successful athletic performance and avoiding injury to most athletes, any intervention modality may benefit by linking excessive substance use to poor athletic performance and/or increased injury risk. For example, in one recent study the researchers provided estimates of how long a single or back to back binge drinking episode might impact their athletic performance (3 days for a single binge; 5 days for back to back binges).

SUMMARY

Researchers have established that for some substances, like alcohol use, certain groups of athletes are more at risk for using the substance than their nonathlete peers. It is fairly well-established that college athletes on an average drink more than students not participating in athletics, and there is evidence to suggest that a similar pattern occurs among high school athletes. For tobacco and recreational drugs, though, evidence suggests that athletes use the substances at rates that are similar to or less than their nonathlete counterparts. In general, little is known about the prevalence rates of substance use among professional or other "elite" athletes.

The reasons that athletes drink alcohol more than others or use illicit drugs in the face of potential sanctions such as suspension and possible loss of income are complex. Certainly athletes are susceptible to the same factors that influence substance use in the general population. It is also possible that certain factors unique to sport, such as the cultural association between alcohol and athletics or a normative team-related environment that promotes alcohol use, in part cause some athletes to drink at high rates.

There currently exist a number of empirically supported treatment and preventive interventions in the substance abuse arena. Although most of these interventions have not been tested specifically among athletes, there is no reason to suspect that they would not be effective among them. Researchers have recently examined the effects of interventions that contain information targeted specifically for athletes, with findings that are promising.

SEE ALSO

Anabolic-androgenic Steroid Use and Dependence, Binge Drinking, Impulsivity, Disinhibition, and Risk Taking in Addiction, Peer Influences on Addiction

Glossary

National Collegiate Athletic Association (NCAA) the largest governing body for college sports in the United States.

Performance-enhancing drugs substances designed to improve athletic performance that are generally considered against the rules of a sport, such as steroids and human growth hormone.

Personalized drinking feedback personalized feedback about one's alcohol-related behaviors that is based on an individual's responses to a series of questionnaires.

Social norms the perception of alcohol use among individuals in a certain reference group (e.g. perceived alcohol use among the typical college student).

Further Reading

Collins, T., Vamplew, W., 2002. Mud, Sweat, and Beers: A Cultural History of Sport and Alcohol. Berg, New York.

Doumas, D., Haustveit, T., Coll, K.M., 2010. Reducing heavy drinking among first year intercollegiate athletes: a randomized controlled trial of web-based normative feedback. Journal of Applied Sport Psychology 22, 247–261.

Dunn, M., Thomas, J.O., Swift, W., Burns, L., 2011. Recreational substance use among elite Australian athletes. Drug and Alcohol Review 30, 63–68.

Goldberg, L., Elliot, D.L., MacKinnon, D.P., et al., 2003. Drug testing athletes to prevent substance abuse: background a pilot study results of the SATURN (Student Athlete Testing Using Random Notification) study. Journal of Adolescent Health 32, 16–25.

Lisha, N.E., Sussman, S., 2010. Relationship of high school and college sports participation with alcohol, tobacco, and illicit drug use: a review. Addictive Behaviors 35, 399–407.

Martens, M.P., Dams-O'Connor, K., Kilmer, J., 2007. Alcohol and drug abuse among athletes: prevalence, etiology, and interventions. In: Tenenbaum, G., Eklund, R.C. (Eds.), Handbook of Sport Psychology, third ed. John Wiley & Sons, Hoboken, NJ, pp. 859–878.

Mays, D., DePadilla, L., Thompson, N.J., Kushner, H.I., Windle, M., 2010. Sports participation and problem alcohol use: a multi-wave national sample of adolescents. American Journal of Preventive Medicine 38, 491–498.

Miller, K.E., Barnes, G.M., Sabo, D., Melnick, M.J., Farrell, M.P., 2002. A comparison of health risk behavior in adolescent users of anabolic-androgenic steroids, by gender and athlete status. Sociology of Sport Journal 19, 385–402.

National Collegiate Athletic Association, 2006. NCAA Study of Substance Use of College Student-Athletes. Author, Indianapolis, IN.

Nelson, T.F., Wechsler, H., 2001. Alcohol and college athletes. Medicine & Science in Sports & Exercise 33, 43–47.

O'Brien, K.S., Kypri, K., 2008. Alcohol industry sponsorship and hazardous drinking among sportspeople. Addiction 103, 1961–1966.

Terry-McElrath, Y.M., O'Malley, P.M., Johnston, L.D., 2011. Exercise and substance use among American youth, 1991–2009. American Journal of Preventive Medicine 40, 530–540.

Walsh, M.M., Hilton, J.F., Ellison, J.A., et al., 2003. Spit (smokeless) tobacco intervention for high school athletes: results after 1 year. Addictive Behaviors 28, 1095–1113.

Wichstrøm, T., Wichstrøm, L., 2009. Does sports participation during adolescence prevent later alcohol, tobacco, and cannabis use? Addiction 104, 138–149.

Yusko, D.A., Buckman, J.F., White, H.R., Pandina, R.J., 2009. Alcohol, tobacco, illicit drugs, and performance enhancers: a comparison of use by college student athletes and nonathletes. Journal of American College Health 57, 281–290.

Relevant Websites

http://oade.nd.edu/educate-yourself-alcohol/alcohol-and-athletes – Office of Alcohol and Drug Education.

http://www.acsm.org/docs/current-comments/alcoholandathleticperformance.pdf – American College of Sports Medicine.

CHAPTER

60

Cultural Influences on Youth Alcohol and Drug Use

Sharon Hsin Hsu

Center for the Study of Health & Risk Behaviors, University of Washington, Seattle, WA, USA

OUTLINE

Recent statistics suggest that ethnic minority population in the US is increasing. For example, according to US Census Bureau, by the year of 2050, it is expected that the Hispanic and Asian populations will triple and non-Hispanic whites (individuals who identify as white but not of Spanish/Hispanic/Latino origin) will represent about 50% of the total population. Figure 60.1 displays a distribution of the current US population by ethnicity and Hispanic origin and projections for 2050. According to a 2003 Census report, ethnic minority youth accounted for a large percentage of the total youth population. For example, Hispanics were the most likely to be preschoolers relative to other racial and ethnic groups; more than 10% (or 4.2 million) of its total population is in this age group. Native Hawaiians and other Pacific islanders reported the highest percentage (18%) of elementary-school age youth. American Indians, Alaska natives, native Hawaiians, and other Pacific Islanders had the highest percentage (8%) of high school age youth. Taken together, with the increase in population, it is imperative for federal and state agencies to allocate resources for service and research dedicated to addressing health-risk behaviors among ethnic minority youth.

In 1993, the National Institute on Drug Abuse (NIDA) composed a report to summarize research advancement, methodological issues, as well as new direction for research in understanding drug abuse by ethnic minority youth. In the introduction statement, De La Rosa and colleagues posited,

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00060-7

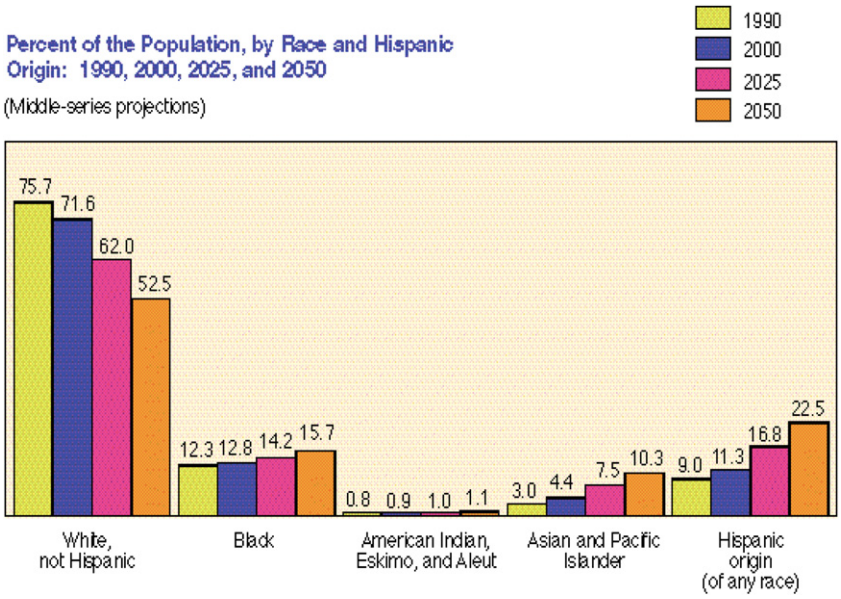


FIGURE 60.1 Percent of the population, by race and Hispanic origin reported by the US Census Bureau.

> *"The limited literature suggests that, because of* cultural influences, *unique economic situations, and formal and informal social network systems, the drug-using behavior of minority youth may vary significantly from that of non-minority youth. There is an urgent need for etiologic research that investigates the interactive roles of intrapersonal, interpersonal, familial, cultural community, and other larger societal factors on the onset, casual use, escalation to use, maintenance, development of dependence, cessation of use, and relapse to use of licit and illicit drugs among minority youth. Studies are also needed that would investigate protective factors among minority children who are at risk but have refrained from using drugs or from escalating to abuse from initial limited exposure (pp. 1)."*

This statement highlights the potentially important role of culture in substance use. In addition, several researchers, such as Castro and colleagues, have indicated that the "culturally blind" approach is not sufficient given that it ignores the service needs of racial/ethnic minority individuals, especially for those who are relatively traditional and/or low in level of acculturation. Thus, Catro and colleagues suggest that the examination of cultural variables may enhance the cultural relevance, efficacy, and effectiveness of prevention and treatment programs designed to serve racial/ethnic minority people.

This chapter will examine recent studies that have tested cultural variables as risk or protective factors among ethnic minority youth in the US Ethnic minority youth are 24 years of age or younger and include Asian American and Pacific Islander youth; Hispanic youth; African American youth; and American Indian and Alaska Native youth. These definitions are adapted from the NIDA monogram published in 1993. A greater proportion of the chapter will be dedicated to Asian American and Pacific Islander youth and Hispanic youth given that there has been a significant advancement in the literature with regard to these two ethnic groups. We will also discuss research on youth outside of the US in a later section.

INCLUDING CULTURAL VARIABLES IN SUBSTANCE USE RESEARCH

According to the US Department of Health and Human Services (DHHS), culture has been defined as a "common heritage or set of beliefs, norms, and values, shared group attributes and a system of shared meaning.". Castro and Hernández-Alarcón suggest that culture is both an environmental and a behavioral variable, and therefore can be divided into two basic types: environmental elements and psychological elements. The environmental elements consist of local environment and cultural creations (i.e. works of art, buildings, and community norms). The psychological elements consist of beliefs, attitudes, expectations, values, and family norms. Specifically, cultural beliefs refer to a historically shaped manner of "viewing the world" and the way in which people "interpret the meaning of objects and events." These beliefs also offer problem-solving strategies, including prescribed approaches of managing stress. Although this conceptualization is helpful, Kitayama and Cohen suggest that it is important to be aware that there are other ways to define culture.

In the context of substance abuse research, Castro and Hernández-Alarcón suggest that cultural variables are generally in the domains of interpersonal relations and

personal traits. The interpersonal relations domain includes variables such as familism, individualism–collectivism, and *tui lien* (loss of face). The personal traits domain includes variables such as acculturation, enculturation, ethnic identity, and biculturalism. For a complete list of variables, please refer to Castro and colleagues' original article. In the current chapter, we will focus on variables that have been tested frequently in the empirical literature. A brief overview of acculturation and cultural identification is provided, as these constructs have received the most attention from the research community.

Acculturation

Snowden and Hines posited that acculturation is broadly defined as the degree to which members of an ethnic group participate in the cultural traditions, values, and practices of the dominant society. This definition has been used in empirical research frequently. Acculturation is frequently measured by multiple indices. Some examples include nativity (i.e. US-born versus foreign-born), the length of time an immigrant has lived in the US, age at immigration, language abilities, and the context in which the language is used (e.g. home). Generally, the US-born individuals are considered to be more acculturated than foreign-born individuals. The longer an immigrant stays in the new country or the younger the age at immigration, the more acculturated they are likely to be. Higher ability to speak English tends to indicate higher level of acculturation. Indeed, Zane and Huh-Kim suggested that speaking English has often been considered to be the strongest acculturation measure linked to different health outcomes. Theoretical models proposed by many scholars have frequently regarded acculturation as a cause of stress and impaired mental and behavioral health among ethnic minorities. Conversely, Wolsko and colleagues pointed out that *enculturation* has been considered as a positive influence through providing a coherent, connected, and grounded identity. It has often been regarded as part of a healthy lifestyle and a buffer for the negative effects of stress.

Cultural Identification

According to Oetting, cultural identification is defined as "a persistent, long-term underlying characteristic that organizes cognitions, emotions, and behaviors." Individuals with high identification with a particular culture tend to recognize themselves as adapted to that culture. They will make evaluative judgments about people and events that are based on cultural beliefs and choose behaviors that are culturally congruent. This in turn will make them successful in participating in activities in such a cultural context. Oetting also posited that cultural identification is a personality trait and the development and maintenance of cultural identification are based on interactions with the environment.

RATES OF SUBSTANCE USE AMONG ETHNIC MINORITY YOUTH IN THE US

The Substance Abuse and Mental Health Substance Administration (SAMHSA) is an excellent resource for information on the latest trends of substance use and abuse among ethnic minority youth in the US. Below is a brief summary of substance use and abuse trends and treatment data by ethnicity. Most of the citations can be found electronically on the SAMHSA website.

Asian American and Pacific Islander Youth

Asian American youth aged 12–17 were less likely to have used alcohol or engaged in binge drinking during the past year, relative to Hispanic, White, or American Indian/Alaska Native (AI/AN) youth. According to the National Institute on Alcohol Abuse and Alcoholism, binge drinking is defined as consuming five or more alcoholic beverages within a 2-hour period. Additionally, a study conducted by SAMHSA indicates that there are subgroup differences in substance use patterns. Filipino youth had the highest rates of past year alcohol use (29.5%), followed by Korean youth (24.9%), Japanese youth (22.2%), Chinese youth (19.6%), and American Indian youth (16%). Based on these figures, Filipino youth were more likely to have reported alcohol consumption during the past year than Chinese or Asian Indian youth. In terms of binge drinking, Another SAMHSA study indicates that Filipino youth reported the highest rates of past year binge drinking (5.8%), followed by Asian Indian youth (5.4%), Japanese youth (4.5%), Korean youth (3.5%), and Chinese youth (1.1%). Additionally, there is a gender difference in alcohol consumption: female Asian youth (19%) were less likely to use alcohol in the past year compared to their male counterparts (26%).

Epidemiology research also suggests that older Asian youth may be at risk for developing alcohol use disorders. For instance, according to Grant and colleagues, the rate of Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) alcohol abuse among Asian females between the ages of 18 and 29 years old increased significantly between 1991/1992 and 2001/2002. Rate of DSM-IV alcohol dependence among Asian males of the same age group also increased significantly over the 10-year period. Young adult Asian males have

been identified as an emerging high-risk group for developing alcohol use disorders.

Based on the most recent available statistics provided by SAMHSA, substance abuse treatment admissions for Asian and Pacific Islander youth aged 12–17 increased by 52% between 1994 and 1999. Most admissions were first treatment episodes (75%). Primary substances of abuse among Asian youth were marijuana, alcohol, and stimulants, with marijuana being the leading substance. Youth were more likely to enroll in treatment through self- or individual referrals or school referrals relative to the total youth treatment population.

Cigarette smoking among Asian youth has received a fair amount of attention in the empirical research literature. For example, Chen and colleague examined patterns of smoking behavior among 7–12 grade Asian American youth ($N = 1810$) in California. Data was extracted from the 1990–1996 California Tobacco Survey and the California Youth Tobacco Survey. Subgroup differences in lifetime smoking prevalence were found, with Filipinos (18.9%) reporting the highest smoking rate, followed by Japanese (17.3%), Koreans (16.3%), other Asian American youth (13.7%), and Chinese (11.0%). A similar pattern of subgroup difference was also found in the 30-day smoking rate. Gender by subgroup analyses revealed that for Chinese and Koreans, smoking prevalence was higher among males (13.4 and 18.8%, respectively) than females (8.3 and 15.9%, respectively). However, for Japanese and other Asians, smoking prevalence was higher among females (18.6 and 19.3%, respectively) than males (14.3 and 13.1%, respectively). Age of smoking onset differed among subgroups, with Japanese males (14.4%) and females (13.6%) reporting the latest onset relative to other subgroups.

Hispanic Youth

Similar to Asian American youth, Hispanic youth in the US are of diverse background. It is estimated that most Hispanic youth are Mexican (71.2%), followed by Puerto Rican (11.8%), Central or South American (10.9%), Cuban (3.0%), and other Hispanic (3.1%); the majority of them (77.8%) were born in the United States. Similar to patterns among Asian American youth, there are differences in patterns of substance use among Hispanic subgroups. With regard to alcohol use, a SAMHSA study suggests that Cuban youth aged 12–17 had the highest rates of past month alcohol consumption (21.2%), followed by Central or South American youth (17.1%), other Hispanic youth (16.8%), Mexican youth (16.5%), and Puerto Rican youth (14.3%). With regard to drug use, a study conducted by SAMHSA suggests that Puerto Rican youth had the highest rates of past month drug use (13.7%), followed by Mexican youth (10.9%), other Hispanic Youth (8.9%), and Central or South American youth (7.2%). Rates of past month alcohol use, binge alcohol use, and past month illicit drug use were similar between Hispanic male and female youth.

According to SAMHSA, among Hispanics of all ages admitted to substance abuse treatment, alcohol, opiates, and marijuana are the most common primary substances of abuse. The majority of these admissions were male (78%). Some subgroup differences in primary substance of abuse were found: While Puerto Rican admissions most commonly reported opiates as the primary substance of abuse, other Hispanic admissions most commonly reported alcohol.

AI/AN Youth

Alcohol use is an important health concern in the AI/AN community. SAMHSA research indicates that AI/AN youth were more likely than youth of other ethnicity/race to have a past year alcohol use disorder (8.5% versus 5.8%). AI/AN individuals of all ages who enter substance abuse treatment were more likely to indicate alcohol as their primary substance of abuse (63%) compared to admissions of other race/ethnicity (42%). Another important concern among AI/AN youth is tobacco use. AI/AN youth (12–17 years old) have the highest rate of current cigarette use when compared to youth of other race/ethnicity. Although available statistics typically do not examine within-group differences, Beauvais suggests that substance use patterns among AI/AN youth may be influenced by study region (e.g. tribal versus non-tribal) and tribal diversity.

African American Youth

Empirical literature indicates that African American youth have lower rates of alcohol consumption relative to Caucasian youth. Nevertheless, according to Turner and Wallace, drinking-related problems are more severe among African American youth, relative to other populations. SAMHSA research suggests that marijuana use is another concern for African American youth. Marijuana was the primary substance of abuse for 66% of all African American youth entering treatment, indicating a higher percentage than youth of other ethnic/racial backgrounds. Males comprised of 82% of all African American youth marijuana admissions. SAMHSA research also indicates that cigarette smoking may also be of concern for African American youth. Among all age groups (i.e. persons aged 12 and older), African Americans between 21 and 25 years old had the highest past month smoking rate (31%). Male youth were more like to smoke cigarettes than female youth in the

following age groups: 16–17, 18–20, and 21–25. Smoking rates were similar between males and females for younger age groups (i.e. persons aged 12–13 and 14–15).

SUMMARY OF RESEARCH REGARDING CULTURAL INFLUENCES ON SUBSTANCE USE IN THE US

Asian American and Pacific Islander Youth

Empirical literature has indicated a positive relationship between acculturation and substance use among Asian American youth. For instance, as mentioned earlier, Chen and colleague examined smoking behavior and its relation to acculturation among 7–12 grade Asian American youth ($N = 1810$) in California. The study includes a diverse sample of Asian American youth; 33% were Filipino, 19% were Chinese, 13% were Korean, 8% were Japanese, and 26% other Asian American. Acculturation status was assessed by English usage, language spoken at home, and age at immigration to the United States. Results indicate that among all youth in the sample, greater levels of acculturation were linked to greater smoking prevalence rates and earlier age of smoking onset.

Although the literature in general indicates a positive relationship between acculturation and substance use, one study found the opposite pattern. Among a sample of Korean American college students attending a large public university ($N = 108$), Hendershot and colleagues found that acculturation serves as a protective rather than risk factor as it predicted decreased drinking. Acculturation was measured by the Suinn-Lew Asian Self-identity Acculturation Scale developed by Suinn and colleagues. The measure assesses different domains of acculturation such as food and language preference and generational status. Findings in this study highlight the importance of a potential within-ethnic effect between acculturation and alcohol use. Additionally, it may be important to consider the cultural norms of substance use in the ethnic versus dominant culture. For instance, according to the World Health Organization (WHO), the annual average consumption for pure alcohol per capita is 8.4 l in the US for people 15 years and older, and this figure is higher for the Republic of Korea (11.8 l) but lower for China (4.2 l). Thus, while entering the US culture (a culture with lower consumption) may be a protective factor for Korean youth (a culture with higher consumption), it is a risk factor for Chinese youth (a culture with lower consumption).

To further explore the relationship between acculturation and substance use, Hahm and colleagues conducted two studies examining moderators and mediators for the relationship between acculturation and alcohol use among Asian American youth. In both studies, the sample was extracted from the National Longitudinal Study of Adolescent Health, containing Asian American adolescents in 9th–12th grade ($N = 714$). About 44% were Filipino, 31% were Chinese, 7.3% were Japanese, 4.9% were Korean, 3.1% were Vietnamese, 2.8% were Asian Indians, and 16.4% were other Asians. In the first study, acculturation was measured by an index of summing the following three variables: English use at home, place of birth, and length of residency in the United States. Findings indicate that acculturation was positively associated with levels of binge drinking. Thus, contact with the US culture for youth from cultures with lower alcohol consumption is a risk factor. Indeed, according to a WHO report, China, Japan, Vietnam, Philippines, and India all have lower annual alcohol consumption per capita than the US South Korean is the only country with higher annual consumption per capita. Serving as the mediator, friends' alcohol and tobacco use explained the relation between acculturation and binge drinking. These findings are consistent with research on the powerful role of peer influence suggested by Oetting and Beauvais. Connection with peers who use substances is the primary direct socialization influence on Asian American youth's involvement in substance use.

In the second study, Hahm and colleagues examined the moderating effect of parental attachment. Participants were divided into four groups based on their acculturation level, ranging from the high to low: (1) English use at home, US-born; (2) English use at home, foreign-born; (3) no English use at home, US-born; and (4) no English use at home, foreign-born. Results indicated that for youth with low parental attachment, the likelihood of drinking was 11 times greater in the highly acculturated group relative to the least acculturated group. However, the likelihood of alcohol use for youth with moderate or high degrees of parental attachment did not differ across levels of acculturation. This finding supports the family interaction theory proposed by Brook and colleagues. Strong parent–child attachment results in the parents having a greater influence on the youth and therefore is a powerful protective factor for preventing Asian youth drug use.

Hahm and colleagues' research demonstrates the powerful role of parental influence on substance use among Asian American youth. Indeed, based on the National Household Survey on Drug Abuse, Asian youth were more likely than youth from other racial/ethnic groups to believe that their parents would strongly disapprove of their consuming one or two drinks nearly daily. Related to this topic, an intriguing empirical question is how parents communicate to youth about substance use within the context of Asian culture. Empirical literature has indicated ethnic

differences in parent communication for other health-risk behavior among adolescents. Markus and Kitayama suggested that this difference may be attributed to how culture shapes parent–child communication. For example, Kim and Ward found that Asian parents tend to use indirect (e.g. implicit and/or nonverbal), rather than direct communication to discuss condom use. Building on prior research, Hsu and colleagues explored potential ethnic differences in parent communication. The study sample ($N = 1{,}104$) was extracted from a gambling study in a large public university. Participants' average age was 19.5 years old (SD = 2.2) and about one-third of them were Asian. Results indicated that, consistent with findings in the prior research, Asian youth reported having less direct communication (e.g. "My parents ask how much I drink"), relative to their Caucasian counterparts.

Another factor that may influence alcohol use and related problems among Asian American youth is Collective Self-Esteem (CSE). Developed by Luhtanen and Crocker, CSE refers to the value one places on being a member of a social group (e.g. gender, ethnicity). CSE contains four domains: membership self-esteem (i.e. judgment of one's worth as a member of their ethnicity), private collective self-esteem (i.e. evaluation of the worth of one's own ethnicity), public collective self-esteem (i.e. evaluation of how others view the worth of one's ethnicity), and importance to identity (i.e. how important being a member of one's ethnicity is to one's self-concept). Hsu and colleagues examined the roles of acculturation and CSE in drinking-related problems among Asian youth ($N = 442$), extracted from the baseline survey for an alcohol intervention study. Participants were from a large public university and the average age was 20.06 years (SD = 1.40). Results indicated that domains of CSE differentially predicted drinking problems among Asian youth. Furthermore, after controlling for alcohol consumption and sex, the relationship between CSE and alcohol-related problems was moderated by acculturation, measured by the S-L ASIA. Findings suggest that lack of adaptation for the worldview and living styles in the US, compounded by lower evaluation of how others view the Asian ethnicity or how well one functions as a member of the Asian ethnicity, strongly predicted alcohol-related problems. It is possible that participants with lower acculturation and lower levels of these CSE factors experienced "marginalization," marginalized individuals are isolated from both their culture of origin and the dominant society. Additionally, findings suggest that lack of adaptation for the worldview and lifestyle in the US, coupled with high regard for Asian ethnicity, also strongly predicted alcohol-related problems. Findings may be attributed to a potential within-ethnic effect not assessed in this study. Some scholars have suggested that for a particular Asian ethnic group (e.g. Korean American), feeling good about belonging to the ethnic group is related to consuming alcohol, if the culture of the ethnic group accepts or promotes drinking.

Hispanic Youth

A recent report based on the National Survey of Drug Use and Health indicates that US-born Hispanic youth were significantly more likely to have used illicit drugs in the past month than foreign-born Hispanic youth. Indeed, this positive association between acculturation and substance use among Hispanic youth has been found by a number of prior studies.

Related to acculturation is stress occurred in the process of adapting to the worldview and way of living in the host society. An earlier study conducted by Vega and colleagues indicated that drinking may be a coping strategy in response to perceived acculturation stress. To further explore this issue, Gil and colleagues examined factors associated with patterns and consequences of early alcohol involvement among male Latino immigrants ($N = 1051$) and US-born individuals ($N = 968$) attending middle schools in South Florida. Youth and their families were mostly Cuban (40%) and Nicaraguan (13%). An important strength of the study was the ability to examine how nativity may influence the longitudinal relations among variables including acculturation, familism, and alcohol involvement. Familism refers the tendency to live close by the family, as well as the use of family network as a source of emotional and social support. Using structural equation modeling, Gil and colleagues found that acculturation and acculturative stress influence alcohol use mainly through "deterioration of Latino family values, attitudes, and familistic behaviors." (pp. 443) For immigrant youth, the process of acculturation decreases problems related to mastery of the English language, which is an essential tool for survival. At the same time, this process of acculturation increases other aspects of acculturative stress (e.g. perceived discrimination; acculturation conflicts within the family). Among youth born in the US, elevated acculturation decreases acculturation stress. At the same time, acculturation among US-born youth also decreases values of familism and parental respect. Taken together, these findings suggest that depending on youth's nativity (US- versus foreign-born), acculturative stress influences alcohol use via different pathways.

In addition to nativity, researchers have used other methods to assess acculturation; one of which is language use and the context in which the language is spoken. In three studies, Epstein and colleagues demonstrated the role of linguistic acculturation and substance use using the same dataset, including Hispanic 6th and 7th graders in 22 New York City schools with low

socioeconomic status. This sample was predominantly Puerto Rican (47%) and Dominican (17%) youth. According to the WHO, both countries have lower annual alcohol consumption per capita for people 15 years and older compared to the US (6.8 and 7.3 l of pure alcohol, respectively), indicating contact with the host culture may be a risk factor. Results confirmed this hypothesis. A greater proportion of youth who spoke both English and Spanish with their friends tried alcohol and consumed two or more drinks per occasion than those who only spoke English with their friends. Additionally, for boys only, a higher proportion of students who spoke English and Spanish with their parents had tried alcohol and had been drunk, than students who spoke only Spanish with their parents.

In another study, Epstein and colleagues expanded the scope of their research by examining the relation between linguistic acculturation and polydrug use (i.e. cigarette smoking, alcohol use, and marijuana). Frequency of marijuana smoking was higher among youth who spoke English with their parents than those who spoke Spanish with their parents. By their second year in middle school, frequency of polydrug use was greater among youth who spoke English with their parents and bilingual students who spoke both English and Spanish with their parents than those who spoke only Spanish with their parents. Taken together, these two studies suggest that bilingual Hispanic youth may be a high-risk group to develop problematic substance use.

In the last study, to investigate the underlying mechanisms of the relationship between linguistic acculturation and substance use, Epstein and colleagues examined the following variables as mediators: perceived peer smoking norms, perceived peer drinking norms, and psychological distress. Results indicate that relative to students who spoke Spanish with their parents, bilingual students perceived that a higher proportion of their peers drank. This higher perception of peer drinking norm was linked to greater polydrug use when linguistic acculturation was controlled. Thus, peer drinking norms mediated the relationship between linguistic acculturation and polydrug use while peer smoking norms and psychological distress were not factors that explain the relation between linguistic acculturation and substance use.

Several scholars have postulated that cultural identification also plays an important role in substance use among Hispanic youth. In a recent study, Casas and colleagues examined the relationship between level of cultural identification and tobacco use among 1672 Mexican American youth (6th, 8th, 10th, and 12th graders) attending public schools located in a predominantly agricultural community in California within 100 miles of the United States–Mexico border. Some students were from migrating families, others reported having parents who were settled migrants and did not migrate across the border or to any other part of the United States. The Orthogonal Cultural Identification Scale (OCIS), developed by Oetting and Beauvais, was used to assess level of cultural identification with the following: (1) Mexican American/Spanish, (2) American Indian, (3) Anglo/White American, and (4) other. Here is a sample item: Some families have special activities or traditions that take place every year at particular times (such as holiday parties, special meals, religious activities, trips, or visits). How many of these special activities did your family have when you were growing up that were based on…culture? Results indicate that relative to youth with high Mexican American/Spanish cultural identity, youth with lower identification were more likely to use cigarettes on the daily basis. Additionally, youth who identified highly with the Anglo/White American culture were more likely to use cigarettes daily. Youth from non-migrant families or who identified with a traditional Mexican American/Spanish culture were more likely to perceive tobacco use on a regular basis as harmful.

AI/AN Youth

As indicated previously, tobacco use is an important health concern in the AI/AN community. A number of studies have investigated the relationship between culture and tobacco use among AI/AN youth and the findings have been mixed. For example, following predictions in the problem behavior theory developed by Jessor and Jessor and the social ecological theory developed by Bronfenbrenner, Yu and colleagues examined how mental health (e.g. conduct disorder) and environmental factors, including family, social, and cultural domains, were associated with American Indian youth tobacco use. A stratified random sample of reservation ($N = 205$) and urban American Indian ($N = 196$) youth residing in a Southwestern area participated in the study. 65% of the reservation youth and 53% of the urban youth reported lifetime tobacco use and this difference was statistically significant. Cultural factors, including cultural pride and engagement in cultural activities, were measured by a modified version of the OCIS. While mental health and environmental factors in the family and social domains were predictive of tobacco use, cultural factors and living setting (reservation versus urban) were not significant predictors of tobacco use in this sample of adolescents.

Although LaMaster and Yu have obtained nonsignificant findings in their studies, a recent study found an interesting pattern between cultural identification and smoking. Wolsko and colleagues surveyed a large sample of Yup'ik residents in the Yukon-Kuskokwim

Delta region of Alaska. Participants ($N = 853$) included both youth and adults, ranging from 14 to 94 years old. The OCIS was used to assess level of identification with White and Yup'ik cultures. Findings indicated that prevalence of smoking was quite high for youth aged 14–19: About 35% reported smoking any type of tobacco and about 40% reported using Smokeless Tobacco (SLT). Younger participants in the study tended to smoke cigarettes more, as well as to use both cigarettes and SLT more. In the overall sample, SLT users identified more with a Yup'ik lifestyle and less with a White lifestyle, speaking their traditional language more frequently, and consuming more traditional food and medicine. This in turn suggests enculturation as a risk factor for SLT, given that the use of SLT is consistent with the traditional culture of Yup'ik. Conversely, cigarette smokers tended to identify less with a Yup'ik lifestyle and reported using drugs and alcohol as means of handling stress. This in turn suggests acculturation as a risk factor for cigarette smoking. Theoretical models proposed by O'Nell and others have frequently described acculturation as a predictor or a cause of profound stress and impaired mental and behavioral health among ethnic minorities. Wolsko and colleagues' findings challenge the existing conceptualization of enculturation as a protective factor for substance use and researchers must carefully evaluate expressions of enculturation before designing intervention and prevention programs.

African American Youth

It has been suggested that African American is typically viewed as a racial group rather than an ethnic group, thus little research has tested the relationship between conventional cultural variables (e.g. acculturation) and substance use among African American youth. Available research has examined the relationship between ethnic identity and substance use among African American youth and some protective effects of ethnic identity were found by several studies. Among high school age youth, Cladwell and colleagues examined the influence of racial identity and parental support on alcohol use among a sample of 488 "academically at risk" African American youth entering high school in a Midwestern city. Academically at risk was defined as having an overall GPA 3.0 or below when youth were in the 8th grade. Two dimensions of racial identity were measured: (1) racial centrality, defined as the significance that one places on race in defining oneself ("Being Black is a major part of my identity."); and (2) private regard, defined as the extent to which one feels positively about Black people ("I am proud of Black people."). These subscales were derived from the Multidimensional Inventory of Black Identity, developed by Sellers and colleagues. Results indicate that private regard and perceived support from a father were associated with less alcohol consumption for these youth after controlling for age and gender. Private regard was also associated with less alcohol consumption for youth who considered race as a more central part of their identity.

Among college age youth, Pugh and Bry examined the influence of ethnic identity on alcohol and marijuana use among Black university students ($N = 167$). Ethnic identity was measured by one subscale in the Multigroup Ethnic Identity Measure-Revised (MEIM-R). A sample item includes, "I have a clear sense of my background and what it means for me." After accounting for year in school, sex, and friends' substance use, ethnic identity score negatively predicted beer/hard liquor use, wine use, and marijuana use, accounting for approximately 31, 6, and 4% of the variance, respectively.

SUBSTANCE ABUSE TREATMENT FOR ETHNIC MINORITY YOUTH

The above review demonstrates the important role of culture in substance use among ethnic minority youth. Given these findings, Castro and Hernández-Alarcón prevention and treatment programs that ignore cultural issues may be ineffective when administered to racial/ethnic minority youth. To address this important public health concern, a number of programs have been developed for ethnic minority populations or a specific ethnic population. In a systematic review of evidence-based treatment for ethnic minority youth, Huey and Polo suggested that Multidimensional Family Therapy (MDFT) developed by Liddle and colleagues, is the only efficacious treatment for drug-abusing ethnic minority youth. MDFT is a multicomponent program that addresses the multiple systems (e.g. family, school, work, peer) that contribute to the development and maintenance of substance use. Therapists aim to: (1) teach youth communication and problem-solving skills; and (2) change negative family interaction patterns. Liddle and colleagues found MDFT led to more rapid reduction in drug use than group-based cognitive behavioral therapy (CBT) among a diverse sample of ethnic minority youth. Additionally, recent treatment outcome studies indicated that for youth with more severe drug use and greater psychiatric comorbidity, MDFT produced better treatment outcomes when compared to individually focused CBT and enhanced services as usual. Liddle and colleagues found that MDFT was more effective than a peer-group intervention in reducing substance use, substance use problems,

delinquency, and internalized distress among young adolescents (average age = 13.73 years)

A number of programs have been developed for youth belonging to a specific ethnic group. We will present two examples here. The Shadow Project, conducted by Boyd-Ball, is a pilot study involving 60 AI families with youth enrolling in substance abuse inpatient treatment. Participants were asked to participate in one of the two treatment modalities: treatment as usual and a brief family-enhanced intervention. In order to adapt the family-based intervention for AI families, culturally relevant assessment and intervention were used. For example, two legends were used to assess the families' skill level in relationship building, problem solving, and communication as a family unit. Preliminary findings were promising; Indian stories were connected to prosocial behavior among youth and the percentage of days abstinent from drug use.

Ma and colleagues have conducted a study to identify cultural variables that enhance the efficacy of smoking cessation curricula for Chinese American youth ranging between 14 and 19 years old. Male youth smokers ($N = 17$) were recruited from community-based Asian American organizations in the Northeast region of the US. Cultural themes (e.g. collective orientation of the Asian culture) were identified and addressed in curricula. Participants joined one of the two treatment groups: a standard smoking cessation curriculum (SC) and a culturally modified program (ACT) for Asian adolescents. Findings indicate a 23.1% quit rate for the SC program and an 18.2% quit rate for the ACT program three months following treatment. For those who continued to smoke, there was a larger decrease in cigarette use among ACT participants than the Standard (SC) group. The ACT protocol has been evaluated in other clinical trials based on a larger sample size of participants.

RESEARCH ON YOUTH IN OTHER COUNTRIES

As indicated earlier, rates for alcohol use in different countries are available for people over 15 years and older in reports produced by the WHO. Although these reports do not focus on youth per se, it provides analyses on alcohol use trends from 1961 to 2006 and a brief review on alcohol policy for each country. Patterns of drinking, drinking-related mortality, and health consequences were also available for some countries. In Europe, the European School Survey Project on Alcohol and other Drugs (ESPAD) provides a comprehensive overview of trends in tobacco, alcohol and substance use among students (aged 15–16 years) between 1995 and 2007. In the most recent report, Hibell and colleagues found that over the 12-year interval, there was a decrease in smoking in a majority of the countries. While trends for alcohol use largely remained the same in the past 12 months and the past 30 days, there was a continuous increase of heavy episodic drinking. Additionally, all countries showed no increase for marijuana use in the past 30 days. With regard to illicit drugs, there was an increase between 1995 and 2003 but the figures in 2007 indicated a decline. In sum, Hibell and colleagues pointed out that the national cultural context for students in different countries most certainly varied, suggesting opportunities to examine the role of cultural variables in substance use in these countries. Indeed, a literature review in the Psychinfo indicates that there is a growing interest in examining cultural variables for youth outside of US. Research was found in countries such as Canada, Lebanon, Chili, Ireland, Finland, Norway, Germany, and Hungary. For instance, Martin and colleagues discussed the process of modifying Strengthening Families Program, a US family-based program preventing substance abuse and behavioral problems, for Germany youth. No studies were found for youth in African countries, as WHO pointed out that data on alcohol consumption among youth from low-income countries is generally limited.

CONCLUSION

Findings in the above review suggest that contact with the dominant culture (i.e. the US culture) is a risk factor for substance use and misuse for some ethnic minority youth whose native countries have lower alcohol consumption per capital compared to the US. It is clear that the relationship between culture and substance use is complicated in part due to the methodologies employed in assessing the cultural variables. This issue has been well documented. Given that acculturation is a risk factor for many ethnic minority youth in the US, it is imperative to identify the timing of delivering prevention programs. For those who are first generation immigrants, the ideal timing may be shortly after individuals gain some fluency in English and begin to assimilate into the dominant culture. For those who are of later generations (and English fluency is not a factor), the timing of delivery should be similar to that for youth in the dominant culture. However, in designing the prevention program, special consideration should be given to the overall acculturation level of the family and parent–child acculturative gap. It is likely that ethnic minority youth would be at high risk when experiencing a combination of risk factors (e.g. family conflict as a result of acculturative gap) as well as conventional risk factors for substance use (e.g. low

socioeconomic status, male, family history of substance abuse).

Another factor that needs to be considered is the diversity within a given ethnic group. Thus, more within-ethnic studies should be conducted to clarify the relationship between cultural variables and substance use patterns among ethnic minority youth. There has been an increased awareness of subgroup differences within an ethnic minority group. For example, as discussed earlier, recognizing that Asian American is not a homogeneous group due to the rich language and cultural diversity existing in this group, much of the recent literature has been dedicated to understanding how substance use patterns as well as predictors of use may differ in Asian subgroups. A report conducted by the US DHHS has pointed out that this movement will in turn address the concern of treating an ethnic group as homogeneous and thus potentially limiting resources for subgroups that are at high risk for developing substance use disorders.

Castro and Hernández-Alarcón suggested that the examination of moderators as well as mediators has been an increasingly important focus in this line of research. This approach is imperative, given that it allows us to understand the specific conditions and the social and biological processes under which such cultural traits may operate as risk or protective factors. Such information provides insight on for whom we should design intervention and what the active ingredients in an intervention should be. For example, based on Epstein and colleagues' research, school-based prevention and intervention should target bilingual Hispanic youth and their peers. Based on evidence in Asian youth literature, poor parent–child communication must be considered and addressed when treating substance abuse disorders among Asian youth. The initial evidence for culturally relevant interventions is encouraging. Several scholars, such as Lau, Castro, and Hwang, have provided guidelines for adapting a specific intervention or for a specific population are available. This line of work will in turn increase the effectiveness of evidence-based treatment when serving ethnic minority youth with substance use issues.

SEE ALSO

Developmental Risk Taking and the Natural History of Alcohol and Drug Use among Youth

List of Abbreviations

AI/AN	American Indian/Alaska Native
CBT	cognitive behavioral therapy
CSE	Collective Self-Esteem
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, 4th edition
ESPAD	European School Survey Project on Alcohol and other Drugs
MDFT	Multidimensional Family Therapy
MEIM-R	Multigroup Ethnic Identity Measure-Revised
NIDA	National Institute on Drug Abuse
OCIS	Orthogonal Cultural Identification Scale
SAMHSA	Substance Abuse and Mental Health Substance Administration
SLT	Smokeless Tobacco
WHO	World Health Organization

Further Reading

Boyd-Ball, A., 2003. A culturally responsive, family-enhanced intervention model. Alcoholism: Clinical and Experimental Research 27 (8), 1356–1360.

Castro, F.G., Hernández-Alarcón, E.H., 2002. Integrating cultural variables into drug abuse prevention and treatment with racial/ethnic minorities. Journal of Drug Issues 32 (3), 783–810.

Chun, K.M., Organista, P.B., Marín, G., 2003. Acculturation: Advances in Theory, Measurement, and Applied Research. American psychological Association, Washington, D. C.

Epstein, J.A., Doyle, M., Botvin, G.J., 2003. A mediational model of the relationship between linguistic acculturation and polydrug use among Hispanic adolescents. Psychological Reports 93, 859–866.

Huey, S.J., Polo, A.J., 2008. Evidence-based psychocial treatments for ethnic minority youth. Journal of Clinical Child and Adolescent Psychology 37 (1), 262–301.

Hwang, W.C., Wood, J.J., 2009. Acculturative family distancing: links with self-reported Symptomatology among Asian Americans and Latinos. Child Psychiatry and Human Development 40, 123–138.

Lau, A.S., 2006. Making the case for selective and directed cultural adaptations of evidence-based treatments: examples from parent training. Clinical Psychology: Science and Practice 13 (4), 295–310.

Liddle, H.A., Rowe, C.L., Dakof, G.A., Henderson, C.E., Greenbaum, P.E., 2009. Multidimensional family therapy for young adolescent substance abuse: twelve-month outcomes of a randomized controlled trial. Journal of Consulting and Clinical Psychology 77 (1), 12–25.

Markus, H.R., Kitayama, S., 1991. Culture and self: implications for cognition, emotion, and motivation. Psychological Review 98, 224–253.

National Institute on Drug Abuse, 1993. Drug Abuse Among Minority Youth: Advances in Research and Methodology. http://archives.drugabuse.gov/pdf/monographs/download130.html.

Oetting, E.R., 1993. Orthogonal cultural identification: theoretical links between cultural identification and substance use. National Institute on Drug Abuse Research Monograph 130, 32–56.

Suinn, R.M., Rickard-Figueroa, K., Lew, S., Vigil, P., 1987. The Suinn-Lew Asian self-identity acculturation scale: an initial report. Education Psychology Measurement 47, 401–407.

U.S. Department of Health and Human Services, 2001a. Mental Health: Culture, Race, and Ethnicity – A Supplement to Mental Health: A Report of the Surgeon General – Executive Summary. U.S. Department of Health and Human Services, Office of the Surgeon General, Rockville, MD.

U.S. Department of Health and Human Services, 2001b. Mental Health: Culture, Race, and Ethnicity—A Supplement to Mental Health: A report of the Surgeon General (DHHS Publication No. SMA 01–3612). U.S. Government Printing Office, Washington, DC. http://www.surgeongeneral.gov/library/mentalhealth/cre/.

Vega, W.A., Gil, A.G., 1998. Drug Use and Ethnicity in Early Adolescence. Plenum Press, New York, NY.

Wolsko, C., Mohatt, G.V., Lardon, C., Burket, R., 2009. Smoking, chewing, and cultural identity: prevalence and correlates of tobacco use among the Yup'ik – the Center for Alaska Native Health Research (CANHR) Study. Cultural Diversity and Ethnic Minority Psychology 15 (2), 165–172.

Relevant Websites

http://www.hbsc.org/index.html – Health Behavior in School-aged Children: A World Health Organization Collaborative Cross-national Study.

http://www.oas.samhsa.gov/race.htm – Substance Abuse and Mental Health Service Administration. Racial and Ethnic Groups: Reports and Data.

http://www.who.int/substance_abuse/publications/global_alcohol_report/en/ – World Health Organization. Global Status Report on Alcohol and Health 2011.

CHAPTER

61

Substance Use and Mental Health Issues on the College Campus

Jason R. Kilmer, Irene M. Geisner

Center for the Study of Health and Risk Behaviors, University of Washington, Seattle, WA, USA

OUTLINE

OVERVIEW OF SUBSTANCE USE TRENDS IN THE UNITED STATES

Alcohol consumption by college students continues to be both highly prevalent and problematic. Research has shown up to 90% of college students in the US drink alcohol, with 25–50% considered "heavy" or "binge" drinkers (four drinks per occasion for women and five drinks per occasion for men). Recent research by the Monitoring the Future team at the University of Michigan indicates that almost four out of five US college students (79.4%) report that they drink alcohol at least once per year, with 65.8% reporting use in the past month. Of those that make the choice to drink, over half report that they have gotten "drunk," (61.5% in the past year, and 42.4% in the past month). The college years are also associated with increased risk for alcohol use disorders; problems typically first appear in the mid-teens, peak in young adulthood, and then gradually decline, with up to 70% of those meeting criteria for dependence doing so before age 25. In 2007, 18.6 million people over the age of 12 endorsed symptoms related to past-year alcohol abuse or dependence, representing 7.5% of the US population. Full-time college students report much higher rates, with 12.5% meeting criteria for alcohol abuse and 8.1% meeting criteria for alcohol dependence.

Alcohol use does not occur in a vacuum, and other substances are being used. Over one-third (36.0%) of US college students report that they have used an illicit drug in the past year, and marijuana is, by far, the most commonly used illicit substance. In fact, the 32.8% of students who report past-year marijuana use exceeds the 29.9% who report past-year tobacco cigarette use. When marijuana is excluded, 16.9% of US college students report past-year use of an illicit substance. Seven drug categories were associated with over 5.0% of college students indicating past-year use, including: Vicodin (8.4%), Adderall (7.9%), Narcotics other than heroin (7.6%), Amphetamines (7.5%), Oxycontin (7.3%), Salvia (5.8%), and Tranquilizers (5.4%). In addition to the use of any one of these substances on its own, it is possible that students could combine substances (for use at the same time), which opens up the possibility of dangerous drug interactions.

Excessive drinking by students is associated with consequences including damaged property, poor class attendance, hangovers, legal problems, and injuries. Researchers estimate over 1700 alcohol-related

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00061-9

unintentional injury deaths annually among US college students, with an additional 599,000 injured because of drinking and 696,000 hit or assaulted by another drinking college student. The American College Health Association (ACHA) indicates that as a consequence of drinking in the past year, 2.6% of students have gotten in trouble with the police, 1.8% physically injured another person, and 10.7% physically injured themselves. Almost one-fourth (22.3%) of students reported that as a consequence of their drinking in the past year they did something they later regretted, and 19.0% reported that they forgot where they were or what they did. An estimated 474 000 students annually have unprotected sex due to alcohol use and 97 000 report being victims of sexual assault or date rape. Alcohol's impact on cognitive functioning can significantly affect ability to give and obtain consent for sex; as a consequence of drinking within the past 12 months, 1.5% of US students endorse having had sex with someone without giving their consent, and 0.3% of students state that they had sex with someone without getting their consent. Clearly, these consequences have great costs both to the individual (in suffering, pain, losses, etc.) and to society (loss of productivity, property damage, violence, etc.). Out of 12 possible challenges or difficulties from which they could select, almost half of US college students (44.2%) indicated that academics represented the issue that had been most "traumatic or very difficult to handle" in the past year. Survey data also indicate academic impairment is a primary consequence of alcohol use, with 22% of over 40 000 students responding to the Core Alcohol and Drug Survey reporting poor test performance, 28% missing a class, 12% with legal problems, 8% causing damage to others' property, and 30% involved in fights or arguments. However, students might not acknowledge or recognize the impact of alcohol use on academics. Out of 31 possible past-year issues or behaviors that could impact academic performance, only 3.9% of college students acknowledged the impact of alcohol use (ranking it 17th of 31 behind homesickness and in front of a chronic health problem or serious illness), and only 1.6% reported that drug use impacted their academic performance (resulting in a rank of 24 out of 31, behind an injury, including a fracture, sprain, strain, or cut, and in front of pregnancy).

OVERVIEW OF MENTAL HEALTH TRENDS IN THE UNITED STATES

The college years are a particularly stressful time for many young adults while cognitive, emotional, and intellectual development continues at a great pace. The developmental tasks confronting college students include achieving emotional independence from family, choosing and preparing for a career, preparing for relationship commitment and family life, developing an ethical system as well as numerous features related to the transition from adolescence to adulthood, referred to by Arnett as "emerging adulthood." These features include exploration of one's identity, instability (including moving, changing jobs, starting and ending relationships, and shifting educational status), being self-focused (as Arnett notes, not in a pejorative way) by making decisions independently, feeling "in between" as one approaches "full" adulthood, and entertaining numerous possibilities for the future. These tasks require students to develop new social roles and modify old ones, which can result in chronic stress and lead to distress. Specifically, college students' stressors in school and achievement domains, interpersonal relationships, finances, sexual relations, and deviant behaviors can pose significant challenges to overcome. Phenomena such as feeling homesick and moving away from an established network of friends can take their toll on new college students and impact psychological adjustment.

Data from the National Epidemiologic Study on Alcohol and Related Conditions (NESARC) highlights the prevalence of mental health issues in college settings in the United States. Data collected from individuals ranging from 19 to 25 years of age who had attended college in the past 12 months were used ($n = 2188$), and 45.8% had any psychiatric diagnosis in the past 12 months. In fact, 20.4% of students met criteria for past-year alcohol use disorders (12.5% alcohol dependence, 7.9% alcohol abuse), followed by 17.7% with any personality disorder (though they noted this was assessed on a lifetime basis), 11.9% with any anxiety disorder, and 10.6% with any mood disorder. Interestingly, findings and percentages for college student diagnosis were not significantly different than for nonstudents of the same age, suggesting that despite additional opportunities and access to education, students remain at risk. Within mood disorders, 7.0% met criteria for major depressive disorder, 3.2% met criteria for bipolar disorder, and 0.8% met criteria for dysthymia; within anxiety disorders, rates were 8.1% for specific phobia, 3.2% for social anxiety disorder, 2.0% for panic disorder, and 1.6% for generalized anxiety disorder. Personality disorders (again, assessed on a lifetime basis) included obsessive–compulsive (8.2%), paranoid (4.9%), antisocial (4.7%), histrionic (3.5%), schizoid (3.3%), avoidant (2.3%), and dependent (0.5%). Data from the ACHA indicate that over two percent of college students received past-year treatment with medication, psychotherapy, both, or some other treatment for five issues: depression (7.6%), anxiety (7.3%), panic attacks (3.1%), attention-deficit hyperactivity disorder (2.6%), and insomnia

(2.3%). Past-year treatment for eating disorders (not assessed in the study by Blanco and colleagues) was at 0.6% for both anorexia and bulimia.

A review of mental health issues endorsed by students highlights the challenges students encounter during the year. A survey from Fall 2009 examining past-year consequences found 80% of students felt exhausted (not from physical activity), 60% felt very sad, 56% felt very lonely, 46% felt things were hopeless, 30% felt so depressed it was difficult to function, and 6% seriously considered suicide. Hopelessness and depressed mood are increasingly reported at university counseling centers, with 45% of those seeking counseling for any reason being diagnosed with depression. Even depressive symptoms (without a major depressive disorder diagnosis) have been shown to relate to reductions in GPA of up to 0.25 compared to nondepressed students, and up to 90% of students presenting to US counseling centers experience academic difficulties due to their mood. In addition, depressed mood relates to lower self-esteem, peer difficulties, poor psychosocial functioning, less energy, higher levels of externalizing and internalizing problems, and may lead to a major depressive episode.

Stress and anxiety also pose challenges. A majority of students (85%) report they have felt overwhelmed by all they had to do at least once in the past year; in fact, over half (52%) have felt overwhelmed at least once in the past 2 weeks. Another 47% felt overwhelming anxiety at least once in the past year. Also related to emotional difficulties, approximately 37% of college students felt overwhelming anger at least once in the past year. Stress has been suggested as a possible cause of depression, as well as a factor leading to relapse into another depression episode. In college students, stress has been related to more headaches, sleep problems and minor illness (colds), hopelessness, and suicidal ideation. Since more stressful life events have been related to poor psychological adjustment, it is not surprising that college students experience more mental health problems than older adults. Even within the college years, freshmen are more depressed than are seniors. Popular media refers to today's college students as the "therapy generation." The common concerns of leaving home, making friends, finding a relationship, doing well in school, and finding a career path are complicated by more divorced families, more lifestyle choices, alienating effects of technology, too much information, and a poor economy. In addition, sleep deprivation and substance use are common college experiences that can act as stressors related to depressed mood in those vulnerable. In fact, the amount of alcohol consumed was a predictor of less nighttime sleep, later sleep schedules, more sleep on weekends compared to weekdays, and greater delays between bedtimes on weekdays and weekends. Daytime sleepiness appears to mediate alcohol's impact on academic difficulties. Collectively, these issues can impact adjustment, development, perseverance, relationships, and success in the college setting.

OVERVIEW OF THE OVERLAP OF MENTAL HEALTH ISSUES AND SUBSTANCE USE

Depression, anxiety, and substance use often co-occur. This overlap is particularly problematic because substance use can increase the likelihood of high-risk behaviors including suicide attempts. The ACHA survey suggests 3.8% of US students have seriously considered suicide in the past year. Additionally, 1.2% of students seriously considered suicide in the past 12 months as a consequence of drinking. For years, researchers and clinicians have observed a relationship between problematic drinking and depressed mood. Research indicates the risk of having major depression is about three times higher for those with an alcohol use disorder (9.6%) than without (3.3%). Similarly, depressed mood has been associated with increased risk of problematic alcohol use and related problems. The odds of meeting criteria for an alcohol use disorder in the context of depression, compared to those without, are roughly four times greater. Data from the mid-1990s indicated for those with major depression, 21% also had an alcohol use disorder (approximately three times greater than the 7% who were not depressed). More recently, results from the National Survey on Drug Use and Health (NSDUH) demonstrated that 17% of adults 18 and older with a past-year major depressive episode also met criteria for past-year alcohol dependence or abuse, compared to 7% of those without.

Numerous studies have found a relationship between greater alcohol use and/or related problems and depression or depressed mood in both community and college samples. Researchers have found a 2.4 odds ratio for college students with alcohol dependence to have a mood or anxiety disorder, and even non-binge drinkers had a 50% increase in risk of mood or anxiety disorder compared to abstainers. Finally, hopelessness, depression, and high-risk drinking have all been identified as risk factors for completed suicide. Similarly, college students with depressed mood have been shown to be more likely than nondepressed peers to use alcohol and other drugs and have more related consequences. Approximately 82% of students with "poor mental health/depression" report drinking alcohol, half report heavy episodic or binge drinking, and these students report increased likelihood of drinking to get drunk compared to those without depression.

There is a debate about how these co-occurring disorders originate (i.e., which comes first, the chicken or the egg?). One hypothesis is that students experience mental health issues and turn to alcohol in order to manage negative affect (i.e., self-medicate). For example, the tension reduction hypothesis suggests that alcohol reduces stress and other negative affect and is consumed for these tension-reducing effects. Consistent with the self-medication hypothesis, higher negative affect is related to increased coping motives for drinking, and drinking to cope has been shown to be related to greater alcohol use and directly related to alcohol problems even when use is held constant. Furthermore, coping motives appear to moderate the relation between stress and drinking in community and college samples. For example, a 3-week study of moods, situations, and drinking in a college sample found weekday drinking, especially at home, appeared motivated by tension-reduction, and was associated with negative experiences and moods. Thus, negative experiences earlier in the day predicted evening alcohol consumption at home and this was strongest in those endorsing stronger coping motives. Recent studies using frequent monitoring designs have supported and helped refine various aspects of the self-medication and tension-reduction models. Specifically, these studies have shown that the comorbidity of alcohol misuse and depressed mood, as well as drinking for primarily coping reasons, is related to increased risk of negative consequences, decreased likelihood of maturing out of alcohol use problems in the future, and more likelihood of developing chronic alcohol problems postcollege.

Conversely, others have proposed that college student drinking is primarily motivated by social reasons, but due to the deleterious effects of heavy drinking and related negative consequences the drinking leads to depressed mood. For example, excessive alcohol use can exacerbate mood symptoms due to sleep disruption, can impact cognitive functioning, which influences individuals' ability to appropriately process negative events, and can interfere with important role obligations like school/academics, etc., which in turn creates stress and impacts students' mood. It is likely that these relationships are actually bidirectional. For example, some researchers found sadness and hostility predicted subsequent drinking behavior, and drinking was associated with relief from sadness and hostility. However, drinking was also related to increased sadness and hostility dependent upon timeframe and social context, illustrating the complex nature of these relationships. It remains an empirical question that has been largely unaddressed in the field whether treating alcohol first, depression first, or both at the same time is most efficacious, with individual differences likely contributing to the determination. Unfortunately, students experiencing these difficulties are often not identified, and when they are, they often receive no services. Thus, an additional challenge involves connecting people who could benefit from clinical services with appropriate care, earlier, to prevent development of a more significant problem in the future.

SCREENING/OUTREACH EFFORTS

Given the well-documented issues and associated risks for college students, there is still much need for development, testing, and broad scale implementation of screening and outreach (i.e., prevention) efforts. For example, slightly under one-half (49%) of college students in the US report receiving any information about depression or anxiety from their college or university, and slightly over one-half (54%) report receiving information about stress reduction. Those with comorbid depressed mood and problematic alcohol use may face barriers to getting help. In a national community sample, 65% of individuals with a substance use disorder and one or more mental health symptom received no treatment. These results are consistent with research demonstrating college students needing services for alcohol or depression often do not seek or receive such services. For example, though 20% of college students in the 2002 NSDUH met criteria for an alcohol use disorder, only 3.9% of students with an alcohol use disorder had received alcohol services in the past year. Further, only 2.4% of full-time college students with an alcohol use disorder perceived a need for treatment.

Utilization of services for depression also fails to match prevalence of those with a need for treatment. Most people with depression receive no treatment, and the Surgeon General has advocated promoting access to treatment as a more significant need than developing better treatments. On US college campuses, out of 103 suicides reported by counseling center directors, only 19% were current or former counseling center clients. Similarly, up to 72% of US college students who screened positive for major depression felt they needed help, but only 36% received services of any kind. Factors related to not accessing help included students being unaware of or unfamiliar with service options, questioning the helpfulness of treatment, and being uncertain about insurance coverage. Additionally, there was less use of services by students who identified as growing up in a "poor family" and by students who identified as Asian or Pacific Islander. Six reasons for not utilizing services were directly identified by at least 20% of the students with positive depression screens who did not receive services, including the belief that stress is normal in a college setting, not perceiving a need for services, believing things would get better in time without

intervention, not having time to get services, thinking no one could understand their problems, and worrying about what others would think. Routine screening in health care settings is one strategy for identifying students who could be having difficulty with mental health or substance use issues. Only 32.5% of health centers on college campuses routinely screen for alcohol problems, and, of those that assess, only 17% use standardized measures. One possible screening measure that has been used in a college setting to facilitate referral to brief interventions is the Alcohol Use Disorders Identification Test (AUDIT). The AUDIT is a 10-item, self-report measure that assesses alcohol use and related consequences. In a sample of college students, a cutoff of 8 results in sensitivity of 82% and specificity of 78% in comparison to diagnostic interview. Scores between 8 and 15 typically warrant a brief intervention using simple advice and patient education materials, and scores between 16 and 20 could also include continued monitoring. A recent study used the AUDIT to screen students when they arrived for their appointment in the health center or the counseling center (even if alcohol use had nothing to do with the presenting issue). Using this strategy, researchers identified 175 students who screened as engaging in "at risk" alcohol use and who attended a Brief Alcohol Screening and Intervention for College Students (BASICS) motivational feedback session. Students who received the intervention demonstrated reduced alcohol consumption, correction in misperceived norms, and increased use of protective behavioral strategies. Even briefer interventions have been tested in college health settings, in which all interventions were delivered by primary care providers. Students at five universities were screened using a health screening survey, and referred students for a brief intervention if they exceeded cutoffs for past 28-day quantity (more than 50 drinks for men, more than 40 drinks for women) or past 28-day heavy drinking episodes involving consumption of five or more drinks (eight or more episodes for men, six or more episodes for women). The 15-min intervention (which used motivational interviewing in its delivery and utilized a manual consisting of 24 intervention strategies) was followed by a follow-up phone call or email in 2 weeks, an in-person 15-minute reinforcement session in 1 month, and a follow-up phone call or email 1 month after the reinforcement session. Results indicated significant reductions in alcohol use and related consequences, and the authors suggested widespread implementation of screening and brief intervention efforts on campuses, including implementation in health settings.

Screening for depression is also important, and research on the properties of the Patient Health Questionnaire-9 (PHQ-9) for major depression has indicated that a PHQ-9 score greater than or equal to 10 has a sensitivity of 88% (i.e., 88% of people with major depression are correctly identified as being at risk by exceeding the cutoff score on the screening measure) and a specificity of 88% (i.e., 88% of people without major depression are correctly identified as being below the cutoff score on the screening measure). Most recently, a cutoff of 10 and endorsement of having a "very" or "extremely difficult" time functioning in academic, social, or occupational realms has been utilized with college students. In the National College Depression Partnership, screening using the PHQ-9 in primary care settings across eight colleges and universities resulted in more than 800 students with clinical depression being identified, enrolled into a tracking registry, and followed to assess outcomes. Within 12 weeks, over half of the students reported improved functioning, which represents greater intervention response than is typically seen in studies with community samples.

In addition to screening, other steps can be taken to reduce barriers to accessing services. Researchers suggest that colleges and universities could consider educational and awareness campaigns regarding mental health services, given that underutilization of services was associated with not being aware of available services or the potential utility of these services, and not knowing what would be covered (or not covered) by insurance. Additionally, with students suggesting that they were unsure if counseling was indicated (due to the perception that stress and struggling are "normal"), other beliefs and targets could also be indicated. It is important to consider that these campaigns can only be successful if counseling centers are prepared to support and handle increased demand.

COUNSELING CENTER IMPRESSIONS

College counseling centers in the United States have done much to keep up with the changing times and needs of their students. Counseling center staff members are aware of the increased demand and greater severity in psychological problems among students presenting for counseling. In the National Survey of Counseling Center Directors sponsored by the American College Counseling Association (ACCA) and published by the International Association of Counseling Services (IACS), over 93% of 302 directors indicated they have seen evidence supporting the widely reported trend that in recent years there has been an increase in students arriving at counseling with serious psychological problems. Directors were asked to compare their perceptions of presenting issues now to 5 years ago, and indicated increases in severe psychological problems (89.4% perceived an increase), medication issues (75.9%), crisis issues requiring an immediate response

(70.6%), learning disabilities (57.7%), self-injury (55.7%), illicit drug use other than alcohol (46.5%), and alcohol problems (45%). When asked to identify administrative issues or concerns that occupy the time of counseling center directors, the four most frequently endorsed concerns were administrative issues related to students with severe psychological problems (75.5%), balancing the varying demands for counselors' services (71.5%), keeping administration at the institution informed while protecting student confidentiality (69.2%), and the growing demand for services without an appropriate increase in resources (66.2%).

These numbers were corroborated through the annual survey of the Association for University and College Counseling Center Directors (AUCCCD), which invites directors from universities outside the United States to participate (though, of the 385 who participated, 97.4% were from the Unites States). Similar to the findings of the IACS survey, 94% of respondents indicated that the number of students with significant psychological problems is a growing concern on their campus. In fact, 71% of directors felt that the numbers of students with severe psychological problems increased on their campus in the past year alone. However, over two-thirds (67.3%) of centers either lost positions or stayed the same size despite these perceived increases in severity.

Research across counseling centers does not, however, suggest an actual change in severity of psychopathology over time. Rather, it is possible that individual schools have seen a change that is not reflected globally. Alternatively, the National Center on Addiction and Substance Abuse (CASA) at Columbia University (2003) suggests that a range of factors could be contributing to the perceived increase, including an actual increase in mental health problems, greater similarity in the college setting to the general population with more inclusive enrollment, greater availability of medication allowing students to attend college who previously would not have done so, lesser stigma allowing for more students to seek help, students discontinuing or interrupting care with a provider, students discontinuing a medication upon starting college, students using substances while taking prescription medications, increased pressure/competitiveness and sleep deprivation contributing to depression, and fewer students taking time off to become stabilized during high levels of stress or mental health problems.

IMPLICATIONS FOR PREVENTION AND INTERVENTION

When a student with both substance use and mental health issues gets involved in counseling, much of what takes place in the counseling session will likely be impacted by the clinical orientation of the provider. This, too, will impact case conceptualization, and will contribute to whether the provider will approach initial treatment of the mental health issue, the substance use issue, or both (i.e., as co-occurring issues). Certainly the clinician will have to take care when considering diagnoses (e.g., many Diagnostic and Statistical Manual of Mental Disorders, 4th edition text revision (DSM-IV-TR) disorders warn that symptoms cannot be better accounted for by use of a substance). Further, the clinician will need to consider the ways in which use of a substance could contribute to, exacerbate, or even cause a mental health issue of relevance to the student. With almost one-fourth of students reporting that they did something they later regretted as a consequence of drinking, it is likely that depression or anxiety could follow something that occurred while intoxicated (again, a matter of the "which comes first?" debate); without adequate coping strategies, a student could turn to substances to cope with these feelings. Considering the context of these issues will likely be important to a positive clinical outcome.

Designing effective prevention and treatment programs for college students has been a top concern of educators, administrators, researchers, and clinicians alike. Because the risks of using alcohol are high, and students perceive excessive drinking to be normative, more traditional treatments (e.g., 12-step programs, medical model programs) are often not relevant or acceptable for this population. In the past several decades, brief treatments have been designed, tested, and successfully implemented on college campuses.

Efficacious prevention efforts for reducing alcohol use or related consequences exist, and many involve interventions delivered to high-risk individuals. These efforts include cognitive-behavioral skills-based approaches, normative feedback interventions, motivational enhancement strategies, and interventions to challenge alcohol expectancies. One approach with a great deal of empirical support involves delivery of Personalized Normative Feedback (PNF). PNF is individually delivered information designed to correct normative misperceptions and contains feedback regarding: (1) personal drinking behavior, (2) one's perceptions of other's drinking behavior (perceived descriptive norm), and (3) other's actual drinking behavior (actual descriptive norm). Presentation of such information has been found to reduce misperceptions of drinking norms among heavy drinkers and to subsequently reduce drinking for periods ranging from 6 months to 2 years, both as a stand-alone intervention and as a component of larger interventions. Recent findings have shown that computerized and web-based PNF alone, without other intervention components or in-person interventions, have effects on drinking of

a comparable magnitude to more intensive interventions. A range of commercial products provide online, web-based alcohol education, including personalized feedback, and vary in the degree to which they have been rigorously evaluated; certainly, as additional research is conducted, these will provide one possibility for reaching college students.

A variety of treatments for depression have been evaluated, including psychological and pharmacological therapies as well as combinations of therapy and medication. Cognitive therapy, cognitive behavioral therapy (CBT), and medications have all been shown to reduce depressed mood and depression. Strategies such as targeting intrusive thoughts and explanatory styles have been successfully incorporated into treatments. In addition, approaches focusing on mindfulness meditation, religion and prayer, social support, and exercise, have all produced improvements with depressed clients. In general, research has suggested treatment is associated with better outcomes than no treatment. Data are less consistent regarding the advantages of any one treatment over another with meta-analytic studies failing to systematically differentiate therapies.

Providing a stepped care approach, moving from a more education/prevention model for those at lower levels of risk, to a brief treatment for those moderately at risk, and so forth has been used with other mental disorders (e.g. generalized anxiety disorder, substance use, eating disorders). This has recently begun to be viewed as an effective and efficient model for delivering targeted intervention for people along the continuum of depression.

Evidence for effectiveness of brief treatments for depression and depressed mood is available. Alternative implementation strategies have also been shown to reliably improve depressive disorders. Telephone administered therapy, computer administered treatment, bibliotherapy, music therapy, and writing brief narratives are less expensive and intensive, yet effective treatments, especially for those at lower risk on the continuum. However, many of these approaches have not been evaluated with college students, and given the high prevalence rates, costly consequences, and the barriers to treatment for this population, new, and even briefer interventions with easier access need to be designed and evaluated.

Recently, randomized controlled trials have supported the efficacy of web-based targeted/indicated prevention as well as "self-help" interventions with minimal or no clinician contact for depression. These therapies are based on CBT and deliver modules often conducted in person (e.g., core beliefs, automatic thoughts) and/or provide some educational content to raise awareness and encourage utilization of other treatment options. Results show moderate to large improvement in symptoms. Other reported benefits include greater satisfaction with the treatment, ability to customize feedback based on personal or tailored information, removal of common barriers, and economic feasibility. Despite these recent advances in computer- or internet-based depression treatment, studies to date are limited by small sample sizes and short follow-ups, thus more research on these approaches is needed. While some computerized and web-based programs have now been developed and tested for a variety of mental health issues including depression and alcohol use, we are not aware of any published clinical trials of web-based interventions addressing both alcohol use and depressed mood (i.e., a comorbid population) in college students. Certainly, education and prevention to raise awareness and reduce barriers to access, screening to promote early identification and/or detection of those needing services, referral to efficacious interventions and treatment, and research into ways to address the overlap of substance use and mental health issues should be part of an overall strategic plan to improve mental health and reduce comorbidity in college campuses.

SEE ALSO

Alcohol Use Disorders, Prenatal Exposure to Alcohol and Illicit Substances, Symptoms and Course: Alcohol Use Disorder in Adulthood

List of Abbreviations

ACHA	American College Health Association
AUD	alcohol use disorder
AUDIT	Alcohol Use Disorders Identification Test
CBT	Cognitive Behavioral Therapy
IACS	International Association of Counseling Services
NSDUH	National Survey on Drug Use and Health
PHQ-9	Patient Health Questionnaire-9
PNF	Personalized Normative Feedback

Further Reading

American College Health Association, 2010. American College Health Association-National College Health Assessment II: Reference Group Data Report Fall 2009. American College Health Association, Baltimore.

Arnett, J.J., 2005. The developmental context of substance use in emerging adulthood. Journal of Drug Issues 35, 235–254.

Blanco, C., Okuda, M., Wright, C., Hasin, D.S., Grant, B.F., Liu, S., Olfson, M., 2008. Mental health of college students and their non-college attending peers: results from the National Epidemiologic Study on Alcohol and Related Conditions. Archives of General Psychiatry 65, 1429–1437.

Chung, H., Klein, M., Greenberg, S., 2010. The National College Depression Partnership: changing how campuses address depression. NASPA Leadership Exchange 7, 16–21.

Cooper, M., Frone, M.R., Russell, M., Mudar, P., 1995. Drinking to regulate positive and negative emotions: a motivational model of alcohol use. Journal of Personality & Social Psychology 69, 990–1005.

Dimeff, L.A., Baer, J.S., Kivlahan, D.R., Marlatt, G.A., 1999. Brief Alcohol Screening and Intervention for College Students (BASICS). The Guilford Press, New York.

Eisenberg, D., Golberstein, E., Gollust, S.E., 2007. Help-seeking and access to mental health care in a university student population. Medical Care 45, 594–601.

Gallagher, R.P., 2009. National Survey of Counseling Center Directors, 2008. In: Monograph Series Number 8R. The International Association of Counseling Services, Alexandria, VA.

Hingson, R.W., Heeren, T., Winter, M., Wechsler, H., 2005. Magnitude of alcohol-related mortality and morbidity among US college students ages 18–24: changes from 1998 to 2001. Annual Review of Public Health 26, 259–279.

Johnston, L.D., O'Malley, P.M., Bachman, J.G., Schulenberg, J.E., 2010. Monitoring the Future National Survey Results on Drug Use, 1975–2009 (NIH Publication No. 10-7584). In: Secondary school students, vol. I. National Institute on Drug Abuse, Bethesda, MD.

Kilmer, J.R., Bailie, S.K., 2012. The impact of college student substance use: Working with students on campus. In White, H.E., Rabiner, D.L. (Eds.), Substance Use in College Students. Guilford Press, New York, pp. 235-252

Larimer, M.E., Cronce, J.M., 2007. Identification, prevention, and treatment revisited: individual-focused college drinking prevention strategies 1999–2006. Addictive Behaviors 32, 2439–2468.

National Center on Addiction and Substance Abuse (CASA) at Columbia University, 2003. Depression, Substance Abuse and College Student Engagement: A Review of the Literature. The National Center on Addiction and Substance Abuse (CASA) at Columbia University, New York.

Ross, V., 2004. Depression, Anxiety, and Alcohol or other drug use among college students. The Higher Education Center for Alcohol and Other Drug Prevention. [On-line], www.higeredcenter.org/.

Swendsen, J.D., Tennen, H., Carney, M.A., Affleck, G., Willard, A., Hromi, A., 2000. Mood and alcohol consumption: an experience sampling test of the self-medication hypothesis. Journal of Abnormal Psychology 109, 198–204.

Relevant Websites

http://www.halfofus.com/ – HalfOfUs.

http://www.ulifeline.org/ – ULifeline.

http://www.dr-bob.org/vpc/ – The Unabridged Student Counseling Virtual Pamphlet Collection.

http://www.iacsinc.org/ – International Association of Counseling Services (IACS).

http://www.collegedrinkingprevention.gov/ – NIAAA's College Drinking Website.

CHAPTER

62

Costs and Consequences (Morbidity and Mortality) Associated with Adolescent and College Drinking and Related Problems

Ralph W. Hingson, Aaron M White

National Institute on Alcohol Abuse and Alcoholism, Bethesda, MD, USA

OUTLINE

This chapter examines (1) trends from 1998 to 2005 in the magnitude of morbidity and mortality associated with college drinking among 18- to 24-year-old students (earlier reports examined data from 1998 through 2001) and (2) interventions established through scientific research to reduce alcohol misuse among college students.

HEAVY EPISODIC DRINKING AND DRIVING UNDER THE INFLUENCE OF ALCOHOL

National surveys conducted by the Substance Abuse and Mental Health Services Administration (SAMHSA) indicate that from 1999 to 2005 the percent of 18- to 24-year-old college students who drank five or more drinks on an occasion in the previous 30 days increased from 41.7 to 45.2%, a significant 8% proportional increase. Among 18- to 24-year-olds not in college, the percent increased from 36.5 to 40.2, a significant proportional 10% increase.

A greater percentage of 18- to 24-year-old college students compared with noncollege respondents drank five or more drinks on an occasion. However, because only one-third of 18- to 24-year-olds are in college, the number not in college who consumed five or more drinks on an occasion in 2005 exceeded the number of college students who did so (7 884 398 versus 4 351 887). From 1999 to 2005, among 18- to 24-year-olds, the

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00062-0

Published by ELSEVIER Inc.

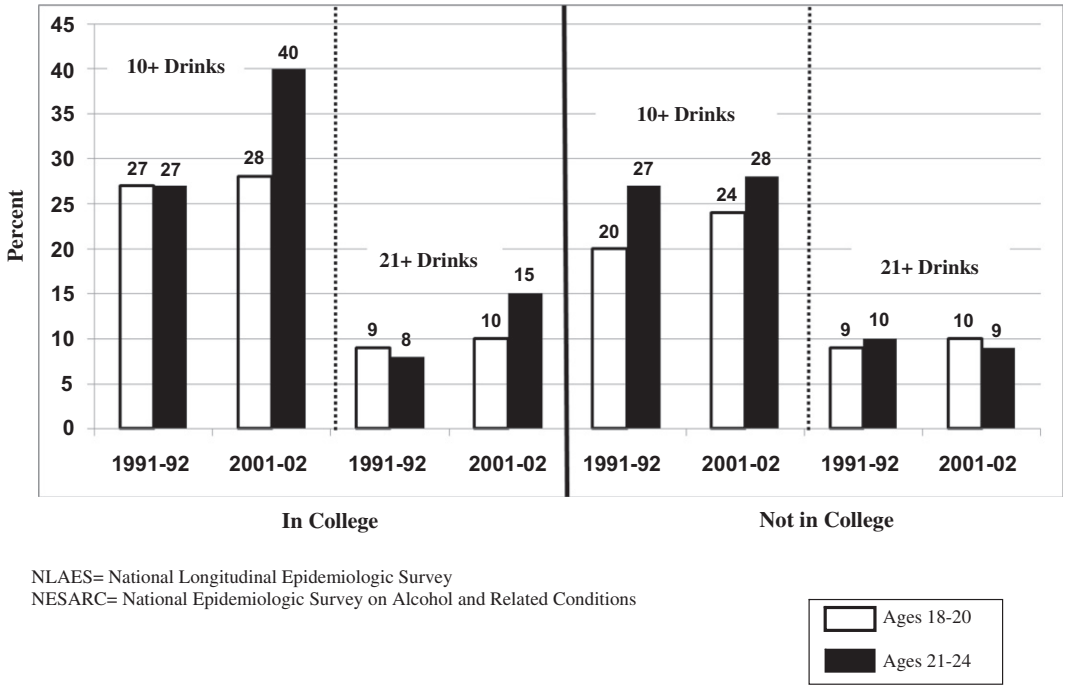


FIGURE 62.1 Consumption of 10+ drinks or 21+ drinks on drinking occasions in the past year by 18–20 and 21–24-year-olds in college versus not in college, 1991–92 (NLAES) and 2001–02 (NESARC).

proportion of college students who drove under the influence of alcohol increased significantly from 26.1 to 29.2%. Among those in the same age group who are not in college, the proportion also increased significantly from 19.8 to 22.8%.

Of note, the increases from 1999 to 2005 in binge drinking and driving under the influence of alcohol occurred among respondents aged 21–24, not those aged 18–20. In each year examined, a greater percentage of 21- to 24-year-olds than 18- to 20-year-olds engaged in these behaviors. Among both 21- to 24-year-olds and 18- to 20-year olds, college students were more likely than same age respondents not enrolled in college to report these behaviors.

Some have argued that the legal drinking age of 21 in the United States drives underage drinking underground into areas lacking adult supervision. Under such circumstances, they maintain persons under age 21 that are more likely to drink very large amounts of alcohol on drinking occasions.

Data collected in two national surveys conducted by the National Institute on Alcohol Abuse and Alcoholism (NIAAA), the 1991–1992 National Longitudinal Epidemiologic Survey (NLAES) ($N = 42\,862$), and the 2001–2002 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) ($N = 43\,093$) offer an opportunity to explore this issue. Both surveys included a question exploring the maximum number of drinks respondents consumed on any drinking day in the previous year as well as questions about respondent age and whether they were college students. Figure 62.1 indicates the percentages of 18- to 20-year-olds and 21- to 24-year-olds who consumed 10 or more or 21 or more drinks on their heaviest drinking days. Twenty-one or more drinks is the equivalent of one-fifth of spirits, a very large and dangerous amount to consume on one occasion. Three findings emerge. First, the highest percentages consuming alcohol at these levels are in college relative to their same age counterparts. Second, a higher percentage of persons ages 21–24 than 18–20 consume these high amounts of alcohol. Third, increases in the percentages consuming these high amounts of alcohol are specifically found in the 21- to 24-year-old age group who are college students.

TOTAL ALCOHOL-RELATED UNINTENTIONAL INJURY DEATHS

Among 18- to 24-year-old college students, deaths from all alcohol-related unintentional injuries, including traffic and other unintentional injuries, increased from 1442 in 1998 to 1825 in 2005, corresponding to increases in rates of death from 18.5 to 19.0%, a 3% increase per 100 000 college students that approached, but did not reach, statistical significance, relative risk 1.03 (95% CI 0.96–1.1). Among all 18- to 24-year-olds, alcohol-related

unintentional injury deaths increased from 4809 in 1998 to 5534 in 2005. Most of the injury deaths resulted from traffic crashes involving alcohol (1357 among college students ages 18–24 and 4114 among all individuals in that age group) in 2005.

Poisoning deaths in the United States increased over 60% among persons of all ages between 1999 and 2004, with the steepest increases involving prescription drug misuse overdoses, particularly opioids. According to the Centers for Disease Control and Prevention (CDC), poisonings claimed the lives of 2290 18- to 25-year-olds in 2005, with 609 involving alcohol, up nearly threefold since 1998. Because the percentages of those who use illicit drugs are similar amongst 18- to 24-year-olds in college and not in college, college students probably account for approximately one-third of the poisoning deaths in that age group.

The role of alcohol in these poisoning death increases warrants close scrutiny. Heavy drinking may impair judgment and contribute to ingestion of larger quantities of drugs, and alcohol may potentiate the adverse effects of drugs.

The National Survey on Drug Use and Health collects data on drug use among college students and others ages 18–25. Similar proportions of college students and others the same age report drug use. Past 30-day use of any illicit drug or prescription drugs used illicitly tended to be slightly higher among persons ages 18–20 than those 21–25 (20% versus 14%). By 2005, those percentages increased to 21.1 and 18.8%, respectively.

Marijuana was by far the most commonly used illicit drug. It was consumed by 18% of college students 18–20 in 1999 and 2005. Among college students ages 21–25, its use rose slightly from 14% in 1999 to 16% in 2005. Use of drugs other than marijuana was reported by 7% of 18- to 20-year-old college students in 1999 and 8% in 2005. Corresponding percentages for college students 21–25 for these years were 5% and 8%.

Not surprisingly, college students and others ages 18–24 who reported binge drinking in the past month were two to three times more likely to report illicit drug and prescription drug misuse than those who did not binge drink.

NIAAA reports have also documented that heavy-drinking college students not only place their own health at risk, they jeopardize the well-being of others. As many as 46% of the 4553 people killed in 2005 in crashes involving 18- to 24-year-old drinking drivers were people other than the drinking drivers. Further, a national survey in 2001 indicated that over 690 000 college students that year Nationwide were hit or assaulted by a drinking college student, and 97 000 students were the victim of a date rape or assault perpetrated by a drinking college student.

INTERVENTIONS TO REDUCE COLLEGE DRINKING

The increase in the past 7 years in alcohol-related traffic and other unintentional injury deaths among 18- to 24-year-olds, both in college and not in college, underscores the need for colleges and their surrounding communities to expand and strengthen interventions demonstrated to reduce excessive drinking among college students and those in the same age group who do not attend college. Numerous individually oriented counseling approaches, environmental interventions, and comprehensive community interventions can reduce drinking and related problems among college students and the college-aged population.

INDIVIDUAL-LEVEL INTERVENTIONS

A meta-analysis of 62 randomized controlled studies of individual-level interventions to reduce college student drinking conducted between 1985 and 2007 with 13 750 participants and 98 intervention conditions reported at short-term follow-up (4–13 weeks post-intervention), and at intermediate follow-up (14–26 weeks post-intervention) intervention, participants reduced the quantity of alcohol consumed and frequency of heavy drinking. At long-term follow-up (27–195 weeks post-intervention), frequency of drinking days and alcohol-related problems were reduced.

Intervention characteristics influenced problem outcomes. Interventions delivered to individuals rather than groups and interventions that used motivational interviewing provided feedback on expectancies or motives, normative comparison, and included decisional balance exercises (e.g. exercises that engage subjects in exploring the pros and cons of particular decisions) were more successful at reducing alcohol-related problems than a range of comparison conditions. In contrast, interventions that used skills training or expectancy challenge components were less successful in reducing alcohol-related problems. They also reported that the magnitude of the effect on drinking diminished over time. In contrast, reduction in alcohol-related problems took longer to emerge but continued in long-term follow-up.

A different review of individually oriented interventions in studies published between 1984 and 2006 found no support for the effectiveness of approaches that only provide information about the health risks linked to alcohol misuse. However, brief motivational interventions (BMIs) received strong support. This approach was found to be effective in reducing drinking problems in 18 of 22 studies that examined that approach,

prompting the conclusion that research continues to strongly support BMIs with personalized feedback delivered individually in groups or as stand-alone feedback with no in-person contact.

Of note, seven studies of mandated populations were reviewed. Mandated populations consist of students instructed to undergo BMIs because they violated alcohol policies. Five of seven studies found that brief motivational feedback interventions were associated with reduced alcohol use or negative consequences. Research conducted in the 2 years since publication of this review supports the effectiveness of BMIs for reducing alcohol consumption as well as the use of other drugs, including computer-based interventions.

An experimental study of college students attending a student health service clinic, found that students screened for heavy episodic drinking who received a two-session brief motivational counseling intervention had significant reductions in typical blood alcohol concentration (BAC), peak BAC, and several other drinking outcome measures at 3 and 6 months follow-up. This is important because most college students at that university went to the student health service at least annually. These findings were recently replicated at five different university health services. These studies are important because routine screening in that setting could have population-wide effects.

NORMATIVE RE-EDUCATION INTERVENTIONS

Research suggests that college students often overestimate the amount of alcohol consumed by fellow students. Misperceptions of normative drinking behavior lead some students to consume more alcohol in an effort to reflect what they perceive to be normal group behavior. A growing body of literature has explored whether informing students of the true norms for alcohol consumption on their campus leads some students to curtail their drinking. This general approach is known as normative re-education or social norms marketing.

A Cochrane review identified 22 randomized trials with 7275 participants and assessed the impact of social norms interventions on college students. It studied the effects of web/computer feedback, individual face-to-face feedback, group face-to-face feedback, mailed feedback, and social marketing campaigns. The web/computer feedback programs achieved significant reductions up to 16 months after the interventions in alcohol problems, peak BACs, frequency of drinking quantity of drinking, and binge drinking. Individual face-to-face feedback produced declines in frequency of drinking at the 6-month follow-up and alcohol-related problems at 4–6 month and 17-month follow-ups. Group face-to-face effects on quantity of drinking and binge drinking lasted only 3 months, and mailed feedback produced no effects. Results from two social marketing studies were inconsistent. Of note, in a recent review summarized earlier in this chapter, four of eight studies examining individual normative feedback found reductions in drinking.

PARENT INITIATIVES

Pre-College Initiatives

An experimental study evaluated two different parent-focused interventions offered to sixth grade students and their parents in 33 rural Midwestern schools. Subjects were randomly assigned to three experimental conditions: the Iowa Strengthening Families Program (ISFP), the preparing for the drug free years (PDFY), and control groups. ISFP sought to improve parent–child relations, strengthening family communication skills, and increase child chopping skills through implementation of a 7-session, 13-h intervention at school with some sessions focused on parents-only and others on children-only. PDFY programs (since relabeled Guiding Good Choices) is a family competency training program offered in five weekly 2-h sessions designed to enhance parent–child interaction and reduce children's risk for early substance initiation. When those sixth graders were reinterviewed during their senior year of high school, compared to students in the control group, those exposed to the ISFP were one-third less likely to report drinking to intoxication. Smaller differences in the same direction were observed among PDFY students.

A recent follow-up study surveyed these subjects at age 21. Compared to controls, the ISFP students reported significantly fewer episodes of drunkenness, frequency of alcohol problems, and cigarette and illicit drug use. Differences between the PDFY and control group students were not significant. Hierarchical latent growth curve models indicated that much of the ISFP intervention's effects resulted from delays in alcohol use initiation during early adolescence. While the study did not report effects specifically among college students, it indicates family interventions targeting sixth graders can delay alcohol use initiation and reduce drunkenness and alcohol problems among college-age youth.

College Initiatives

An experimental design tested the effects of sending parents a 45-page handbook for talking with college

students about alcohol. Parents in the comparison group received a brochure detailing university alcohol policies and consequences of alcohol policy violations.

Of 347 parents in the intervention group, 72% evaluated the handbook and 83% said they had read most or all of it. Students who did not drink prior to college whose parents reviewed the handbook were less likely to start drinking, and those already drinking were less likely to show growth in drinking over the freshman year. This latter finding resulted from effects on female, but not male, students. Other experimental studies found that this parental intervention, in combination with a brief motivational intervention, produced lower levels of alcohol consumption and high-risk drinking among college students compared with a control group. These new findings are important because they indicate parental influence can extend into college-aged youth.

ENVIRONMENTAL INTERVENTIONS: LEGAL DRINKING AGE OF 21

College University Alcohol Policies

In 1998, the Massachusetts Board of Higher Education adopted more restrictive alcohol policies for all schools under its authority, including the following:

- Restricting alcohol to specific, supervised locations
- Required advanced registration of all social events involving alcohol
- Restricting legal possession of alcohol to separate residence halls for students ages 21+
- Mandatory provision of alcohol education and prevention programs
- Established procedures for enforcement of all federal, state, local, and campus regulations
- Requirement for colleges to work with local towns to enforce alcohol laws
- New sanctions for student alcohol violations up to and including expulsion
- Parental notification of alcohol policy violations

In 1999 and 2001, over 1000 randomly selected students at 11 schools each year were surveyed, as were 11 deans or college vice presidents concerning student drinking practices, 14 alcohol policies, and the level of enforcement. Past 30-day drinking declined from 73 to 63% among students under age 21 and from 81 to 74% among campus residents. Heavy episodic past 30-day drinking declined from 56 to 46% among underage drinkers and 37 to 25% among those living on campus. No changes on either measure were recorded for legal age students and those living off campus. No changes were observed in the proportions of students driving under the influence of alcohol or riding with drinking drivers. Declines in college drinking measures over time were greatest in schools where deans cited high alcohol policy enforcement at baseline, and during both surveys. No attempt was made to separate analytically the effects of specific alcohol policies.

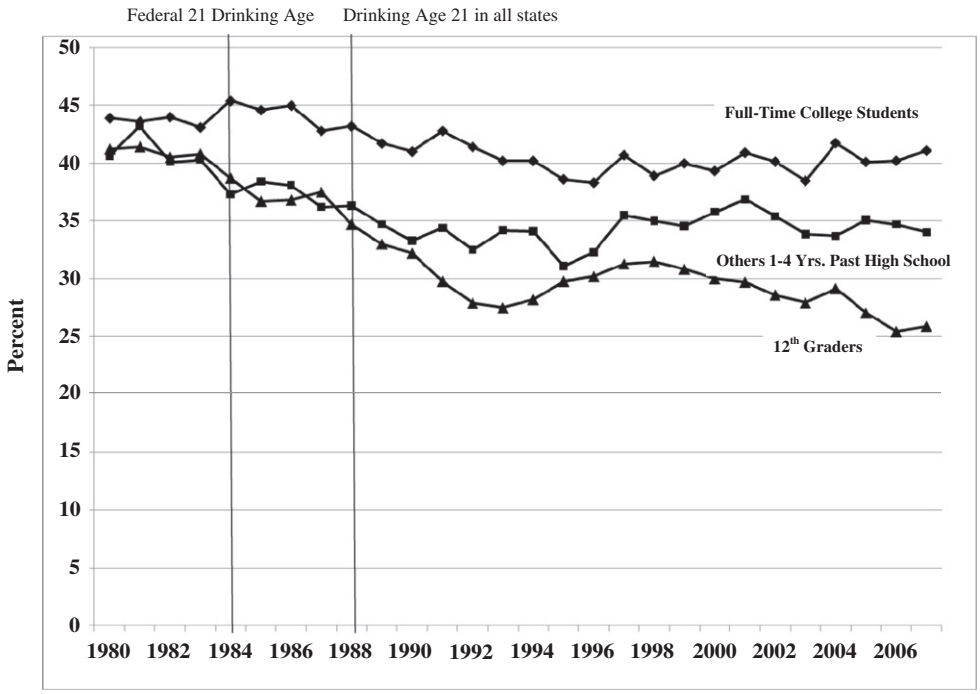
Minimum Legal Drinking Age of 21

In 1984, when 17 states had a legal drinking age of 21, the US Congress passed legislation that would withhold highway construction funding for states that did not make it illegal to sell alcohol to people younger than age 21. By 1988, all states adopted the law.

However, there are some important exceptions. In 24 states, individuals under 21 can possess alcohol with parental or guardian consent and/or presence. In 31 states, parents can legally furnish alcohol to their children who are under 21. Only 31 states and the District of Columbia explicitly prohibit consumption by a person under 21. In 47 states, people under 21 can serve alcohol.

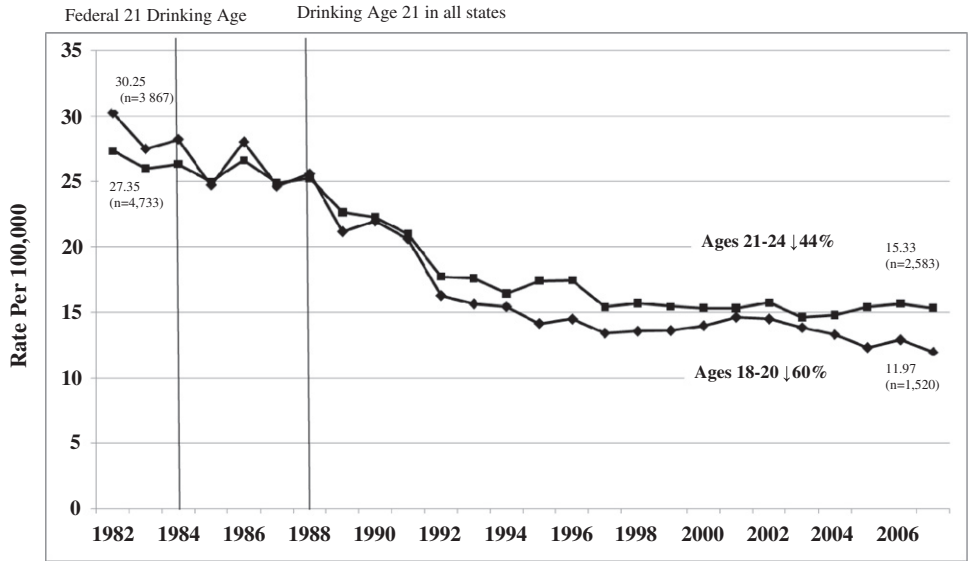
In August 2008, a group of 130 college presidents called for a debate about whether the drinking age should be lowered to age 18. Some suggested, after receiving education about safe drinking levels, that 18-year-olds should be given drinking licenses that would be rescinded if their drinking posed dangers to themselves or others. Given this widely publicized challenge to the legal drinking age of 21, it is worth reviewing evidence on the topic. Figure 62.2 examines trends in the frequency of binge drinking from 1982 to 2007 (five or more drinks on an occasion) from Monitoring the Future, a yearly survey assessment of the attitudes, behaviors, and values of nearly 50 000 8th, 10th, and 12th graders. According to the survey data, binge drinking among high-school seniors dropped from 40% to just over 25%. Among individuals 1 to 4 years past high school, the declines were less, from 40% to just under 35%. Little change was seen among full-time college students. Figure 62.3 examines trends in alcohol-related traffic fatalities among individuals aged 18–20 targeted by the drinking age changes and those aged 21–24 not targeted. Both groups experienced proportional declines, but the declines were greater in the 18- to 20-year age group than in the 21- to 24-year age group (60% versus 44%).

A review of 49 studies of the legal drinking age changes revealed that in the 1970s and 1980s, when many states lowered the drinking age, alcohol-related traffic crashes among people younger than 21 increased 10%. In contrast, when states increased the legal drinking age to 21, alcohol-related crashes among people younger than 21 decreased 16%. A separate review examined 48 studies of the effects of drinking-age changes on drinking and 57 studies on traffic



Source: Monitoring the Future, 2007

FIGURE 62.2 Alcohol: trends in 2-week prevalence of five or more drinks in a row among college students versus others 1–4 years beyond high school. *Adapted from Hingson, R., White, A. Magnitude and Prevention of College Alcohol and Drug Misuse: U.S. College Students Ages 18–24. In Mental Health in the College Community. Kay, J., Schwartz, V. (Eds.). London, United Kingdom: John Wiley & Sons, 2010.*



Source: U.S. Fatality Analysis Reporting System, 2009; U.S. Census Bureau, 2009

FIGURE 62.3 Alcohol-related traffic fatalities, rate per 100 000, ages 18–20 versus 21–24, United States, 1982–2007. *Adapted from Hingson, R., White, A. Magnitude and Prevention of College Alcohol and Drug Misuse: U.S. College Students Ages 18–24. In Mental Health in the College Community. Kay, J., Schwartz, V. (Eds.). London, United Kingdom: John Wiley & Sons, 2010.*

crashes. The study concluded that increases in the legal age of alcohol purchase and consumption have been the most successful interventions to date in reducing drinking and alcohol-related crashes among people under 21.

One recent analysis found significant declines in traffic fatalities among individuals under 21 in states that changed the minimum legal drinking age to 21 prior to the 1984 Federal mandate to raise the drinking age to 21. However, in states that raised the drinking age after

the Federal legislation, the minimum legal drinking age increases were not associated with significant declines in traffic deaths. The analyses controlled for whether states had a seatbelt law, the legal blood alcohol limit, beer taxes, and vehicle miles traveled. Of note, this study did not explore whether the traffic deaths were alcohol related. After adjusting for changes in the population for that age during the time period 1982 to 2007, alcohol-related traffic fatalities among people aged 16–20 declined 62%, whereas those that did not involve alcohol increased 22% (see Fig. 62.4).

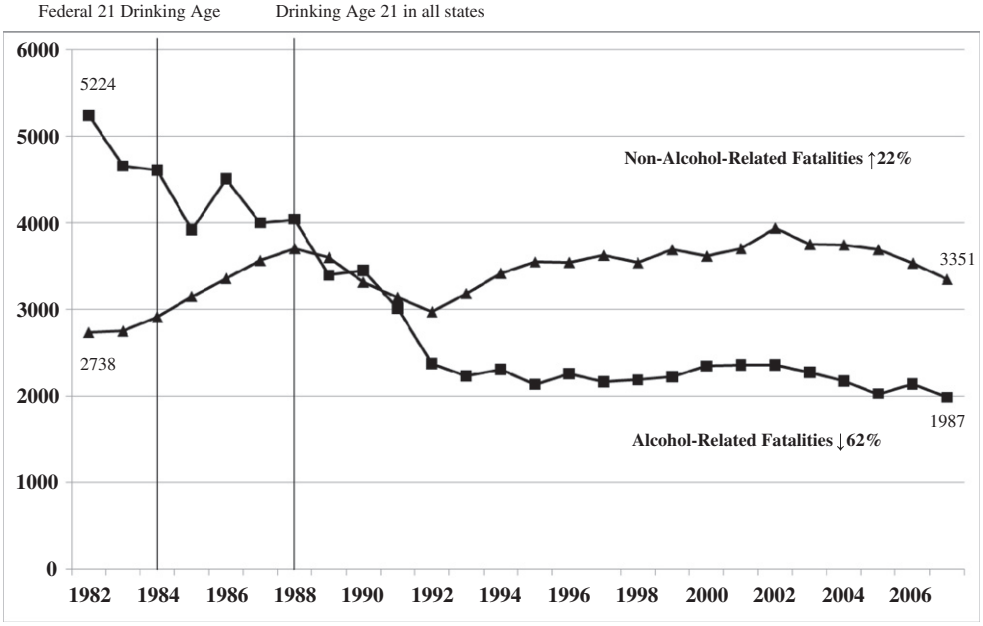
A different study examined trends in the ratio of drinking to nondrinking drivers in fatal crashes in each state annually from 1982 to 2004. This analysis controlled for zero-tolerance laws, graduated license night restrictions, and use/lose laws that target drivers under 21 and could influence their involvement in alcohol-related crashes. It also controlled for 0.10 and 0.08% BAC, per se, legal limits, mandatory seatbelt laws, per capita beer consumption, unemployment rates, vehicle miles traveled, frequency of sobriety checkpoints, number of licensed drivers, and the ratio of drinking to nondrinking drivers aged 26 or older in fatal crashes.

The findings are quite informative. Adoption of the minimum legal drinking age of 21 was associated with a 16% decline in the ratio of drinking to nondrinking drivers in fatal crashes involving those under 21, even after controlling for all the other factors listed above. Of note, other laws targeting drivers under 21 independently predicted lower involvement of drinking drivers in fatal crashes. Use/lose laws and zero-tolerance laws were each associated with 5% declines. Further, laws aimed at adult drivers also independently contributed to declines in the ratio of drinking to nondrinking drivers in fatal crashes: 0.08% BAC laws were independently associated with an 8% decline, 0.10 BAC percent laws a 7% decline, administrative license revocation a 5% decline, and seatbelt laws a 3% decline. Thus, the preponderance of evidence suggests that raising the drinking age to 21 reduced alcohol involvement in fatal crashes involving drivers under 21 and that other laws aimed at drivers of all ages can also reduce alcohol-related fatal crashes involving drivers under the age of 21.

Of note, a recent analysis of over 33 000 adult respondents in two national surveys 10 years apart compared respondents who grew up in states where they legally were allowed to drink prior to age 21 with respondents who grew up in states where the legal drinking age was 21. The analysis, which controlled for numerous potential confounding variables, found that those allowed legally to drink prior to age 21 were more likely as adults to meet both alcohol and drug use disorder criteria.

Zero-Tolerance Laws

Zero-tolerance laws, which make it illegal in every state for those under the age of 21 to drive after any drinking, also have contributed to declines in alcohol-related traffic deaths among people younger than 21.



Source: U.S. Fatality Analysis Reporting System, 2009

FIGURE 62.4 Trends in alcohol related and nonalcohol related traffic fatalities, persons ages 16–20, United States, 1982–2007. *Adapted from Hingson, R., White, A. Magnitude and Prevention of College Alcohol and Drug Misuse: U.S. College Students Ages 18–24. In Mental Health in the College Community. Kay, J., Schwartz, V. (Eds.). London, United Kingdom: John Wiley & Sons, 2010.*

Unfortunately, despite their demonstrated benefits, legal drinking age and zero-tolerance laws generally have not been vigorously enforced. Young drivers are substantially underrepresented in the driving-while-intoxicated (DWI) arrest population relative to their contribution to the alcohol-related crash problem. Stepped-up enforcement of alcohol purchase laws aimed at sellers and buyers can also help reduce alcohol misuse and related problems.

Price of Alcohol

The majority of published studies have reported an inverse relation between the tax on or price of alcohol and alcohol misuse and related negative health outcomes. A National Academy of Sciences review recommended that Congress and State legislatures raise excise taxes to reduce underage alcohol consumption and to raise additional revenues to reduce underage drinking problems. Further, research is needed about the effects of price increases on (1) college students relative to others in the same age and (2) college-age people relative to older people.

Three recent extensive literature reviews examined the relation of alcohol price and tax with consumption and related harms. One analysis of 1003 separate estimates from 112 studies reported very strong evidence that price affects drinking of all types of beverages and across the population of drinkers, from lightest to heavy drinkers.

A separate review of 78 alcohol tax studies meeting the Centers for Disease Control and Prevention's Community Guide inclusion criteria also found consistent evidence across different countries, time periods, study designs, analytic approaches, and outcomes that higher alcohol prices and taxes are associated with reductions in both excessive alcohol consumption and related subsequent harms.

An international review found that when income, the price of other goods, and other factors were analytically controlled for, a rise in alcohol prices was linked to less alcohol consumption and less alcohol-related harm, and vice versa. The authors concluded that increased alcohol prices delay the time when young people start to drink, slow their progression toward drinking large amounts, and on balance reduces their heavy alcohol consumption. Although very high prices for alcohol might stimulate illegal production, in the United States alcohol prices have not kept pace with inflation over the past 60 years.

Alcohol Outlet Density

Higher alcohol outlet density has been associated with increased alcohol-related problems in both cross-sectional and prospective studies, and reducing outlet density may, in turn, reduce drinking-related problems. Prospective research is needed to specifically test whether reducing outlet density will reduce consumption, related problems, and specific effects on college students. One recent study found that higher alcohol outlet density near colleges was related to higher campus rape offense rates.

Comprehensive Community Interventions

Several community-based initiatives have successfully reduced drinking- and/or alcohol-related problems among young people. These programs typically coordinate efforts from the following:

- City officials from multiple departments of city government, school, health, police, and alcohol beverage control, etc.
- Concerned private citizens and their organizations and students, parents, and merchants who sell alcohol.
- Often multiple intervention strategies are incorporated into the programs, including school-based programs involving students, peer leaders, and parents; media advocacy; community organizing and mobilization; environmental policy change to reduce alcohol availability to youth; and heightened enforcement of laws regulating sales and distribution of alcohol and laws to reduce alcohol-related traffic injuries and deaths.

Six comprehensive community programs have achieved reductions in alcohol problems among college-aged youth:

- Communities Mobilizing for Change Program – This program attempted to reduce the flow of alcohol to youth from illegal sales by retail establishments and from the provision of alcohol to youth by adults in the community
- Community Trials Program – Communities formed coalitions aimed at reducing illegal sales of alcohol to youth, implementing responsible beverage service and decreasing drunk driving offenses by increasing awareness of consequences.
- Saving Lives Program – This program tried to reduce drunk driving and related consequences through media campaigns, police training, high-school peer-led education, college prevention programs, increased alcohol outlet surveillance, and other measures.
- Fighting Back Program – This program tried to reduce availability of alcohol through environmental policies and expanded substance abuse screening, counseling, and treatment.
- Sacramento Neighborhood Alcohol Prevention Project – This program included community mobilization, a public awareness campaign, responsible beverage service training, legal drinking

age enforcement, and intoxicated patron law enforcement.
- Reduce Underage Drinking through State Coalitions – This initiative used a community coalition model at the state level in 10 states to mobilize citizens, increase media coverage, and implement policy changes such as alcohol price and tax changes and greater restrictions on commercial and social access to alcohol.

Six studies have now explored elements of the comprehensive community-organizing model as a method for reducing drinking or alcohol-related harms specifically among college students. One study adapted some of the community trials interventions to a college setting. At an experimental university, there was a marked increase in DUI enforcement coupled with a media campaign. The prevention campaign featured DUI checkpoints, media coverage, and a student-designed social marketing campaign designed to increase student perception of risk of arrest for DUI. DUI checkpoints were operated jointly by campus and local city police. Telephone surveys of randomly selected students revealed significant declines in self-reported DUI from pre- to posttest.

The Matter of Degree college–community partnership implemented environmentally based interventions to reduce drinking and related problem behaviors among college students. Interventions included keg registration, mandatory responsible beverage service, campus–community police collaboration on increased wild-party enforcement, substance-free residence halls, and a variety of media efforts. Significant reductions were achieved in binge and frequent binge drinking, frequent intoxication, driving after drinking, alcohol-related injury, and a variety of other alcohol-related problems.

A study of a comprehensive community program focused on underage drinking and driving among 16- to 24-year-olds in Huntington, West Virginia, home of Marshall University (enrollment 18 000 students). Morgantown, West Virginia, home to West Virginia University, with 40 000 students, was selected for comparison. During late winter 2006 and spring 2007, local, university, and state enforcement agencies increased enforcement of drunken and driving laws, including zero-tolerance laws, through low-manpower sobriety checkpoints, saturation patrols, and stepped-up DUI directed patrols. The State Alcohol Beverage Control Administration, with assistance from local and state law enforcement agencies, increased enforcement of the minimum legal drinking age laws. This included enforcement of laws aimed at servers/sellers and underage people, including use of fake identifications. A multimedia campaign that included paid and earned print and broadcast media publicized these efforts.

Roadside surveys of nighttime drivers conducted alcohol breath tests during the fall of 2006, spring of 2007, and fall of 2007. Compliance check survey of underage alcohol purchase attempts produced declines in successful buy attempts, from 43 to 18% in the intervention city. Little change occurred in the comparison city. Reductions in BACs at the roadside surveys in the intervention city showed marked declines in the proportions of drivers aged 16–20, 21–24, and 25 or older at 0.02, 0.05, and 0.08%, respectively. Little change was found in the comparison city.

An examination of a college–community partnership at Western Washington University and another at Washington State University. Police patrols focused on off-campus student parties, and compliance check surveys were used to restrict sales of alcohol to minors. Public forums brought community residents, students, and police together for dialogues about disruptive parties and other neighborhood issues. The colleges also offered alcohol-free late-night activities. Significant restrictions in the prevalence of heavy drinking (five or more consecutive drinks) were observed at the two intervention colleges relative to a comparison college.

An evaluation of the Common Ground Program, a University of Rhode Island/community partnership featuring increased DWI and minimum legal drinking age enforcement, a media campaign, and a safe-rides program. Although the program resulted in increased student awareness of alcohol control measures and a greater perceived likelihood of apprehension for underage drinking, as well as a reduction in police-reported alcohol-related incidents, no changes were observed or reported regarding alcohol use or alcohol-impaired driving by college students.

The largest college/community interventions study to date explored a college/community risk management approach to reduce college student drinking to intoxication and intoxication-related harm and compared seven California public university system colleges/universities to seven comparison colleges/universities in the same state system. The intervention included a social host Safe Party campaign, compliance checks, DUI checkpoints, party patrols, and passage of a social host response cost ordinance. Each college/university focused on at most two settings, focused on the beginning of the school year, underwent a highly specific planning and implementation process that focused more on implementation of intervention rather than motivating colleges. Student surveys during 2 years prior to the intervention were compared with repeat cross-sectional surveys collected twice annually after the intervention. The intervention achieved significantly lower levels of drinking to intoxication at off-campus parties, bars, restaurants, and across all settings studied. They estimated their intervention, on average at each intervention college/university,

produced 1500 fewer students drinking to intoxication and 10 000 fewer incidents of drinking to intoxication. Further analyses revealed no displacement of alcohol problems to other college/community settings. Of note, colleges varied in the extent and intensity with which the interventions were implemented. However, the reported results included even those schools with minimal intervention efforts even though their program effects were negligible.

Taken together, these studies underscore the potential for comprehensive community and college collaborative interventions to specifically reduce alcohol misuse and problems among the difficult-to-reach college student population. Key questions about this approach that warrant future investigation include the following:

- Will a combination of (a) environmental interventions to reduce alcohol availability and enforce alcohol policies, such as minimum legal drinking ages, drinking-and-driving laws, and (b) efforts to expand screening and brief interventions or other individually oriented counseling approaches achieve greater problem reductions than either alone?
- Are programs that target people of all ages more effective in reducing college student alcohol problems than those that focus only on college students?
- Will programs that reduce alcohol misuse among college students and college-aged individuals produce carry-over benefits into adult life?
- Will programs that delay the onset of alcohol use among individuals before they reach their college years reduce drinking and related problems among people in college?
- How can comprehensive campus–community collaborations be sustained over time?
- Which college–community interventions are most effective in reducing alcohol misuse and related problems with the least cost?

CONCLUSIONS

It is ironic that binge drinking and driving under the influence of alcohol continued to rise, and unintentional injuries attributable to alcohol did not decline during a period of time when there was a considerable expansion of the scientific literature and knowledge base regarding how to reduce drinking and related harms among college students. An important research question is how to translate our new knowledge into reductions in alcohol misuse and related problems in the future.

A unique national survey of deans and vice presidents at 351 four-year colleges in the United States in 2008, explored their awareness of a major NIAAA report in 2002 on college student alcohol use and related problems and recommended interventions to reduce student alcohol use. They found nearly 80% of college deans and vice presidents were aware of the report and its recommendations and 77% of their colleges had implemented at least one of the research-supported strategies recommended in the report.

However, the types of NIAAA-recommendations most likely to be implemented were interventions designed to identify individual students with drinking practices that pose risk to health and then offer some form of individual counseling. Two of every three colleges (67%) provided intervention programs for college students who are problem drinkers or at high risk for experiencing drinking-related problems. Most colleges (76%) offered at least one empirically supported program: 66% offered norms clarification, 57% cognitive behavioral skills training, 62% brief motivational interviewing, and 38% expectancy challenge programs. Reviews of these types of interventions since the NIAAA report have reinforced their beneficial value as more new studies have tested and supported their efficacy. Indeed, a promising new finding since the NIAAA report is that these types of interventions can be effective not only with students who voluntarily seek them out but also with students who are mandated to receive them because of alcohol infractions or other disciplinary action.

An important concern, however, is that only a small fraction of college students who might benefit from these programs actually are exposed to them. Based on the NESARC, NIAAA estimated that nearly one in five (19%) college students met alcohol use disorder criteria (dependence or abuse) during the year preceding the 2001–2002 survey, but only 5% of them received any treatment or counseling for their problem. Further, national surveys indicate most adolescent and young adult patients are not asked by their physicians about their drinking and offered advice during medical care encounters about what drinking patterns pose risk to their health.

As noted earlier in this chapter, recent studies have reported efficacy of screening and brief motivational counseling interventions for alcohol in university health services. If student health services were to adopt alcohol screening and brief counseling interventions for at-risk drinkers as a routine part of each university health service visit, that would greatly enhance the potential of these interventions with demonstrated effectiveness to reach a broader segment of the college community and ultimately achieve population-level benefits.

In addition, it is noteworthy that colleges were much less likely to employ environmental policy and comprehensive college/community interventions that have been shown to be effective in reducing youth alcohol problems. Only one in three reported compliance checks to prevent underage drinking were done in their

community, usually (60%) without university involvement. Most administrators had not discussed, planned, or implemented efforts to restrict the numbers of nearby alcohol outlets (79%), increase the price of alcohol (86%), or institute mandatory beverage service training (73%).

Evidence concerning the effectiveness of environmental alcohol policies has expanded considerably since the release of the NIAAA report. As documented in this chapter, a large and growing body of literature supports the beneficial effects of establishing a minimum legal drinking age of 21 in reducing underage drinking and alcohol-related crashes. Further, the majority of published studies have reported an inverse relation between the price of alcohol and alcohol misuse and related negative health outcomes. Higher alcohol outlet density has been associated with increased alcohol-related problems in both cross-sectional and prospective studies, and reducing outlet density may, in turn, reduce drinking-related problems. Prospective research is needed to specifically test whether reducing outlet density will reduce consumption, related problems, and specific effects on college students.

Finally, as also documented in this chapter, several community-based initiatives have successfully reduced drinking and/or alcohol-related problems among young people. These programs typically coordinate efforts of:

- City officials from multiple departments of city government, school, health, police, and alcohol beverage control, etc.
- Concerned private citizens and their organizations and students, parents, and merchants who sell alcohol.

Often multiple intervention strategies are incorporated into the programs, including school-based programs involving students, peer leaders, and parents; media advocacy; community organizing and mobilization; environmental policy change to reduce alcohol availability to youth; and heightened enforcement of laws regulating sales and distribution of alcohol and laws to reduce alcohol-related traffic injuries and deaths. Six of these comprehensive community programs have achieved reductions in alcohol problems among college-aged youth.

In addition, six studies completed since the initial NIAAA report have now explored elements of the comprehensive community organizing model as a method for reducing drinking or alcohol-related harm, specifically among college students. Most achieved reductions in college student drinking or related harms.

The growing evidence on effectiveness of alcohol policy and comprehensive community interventions should prompt more colleges to work together with their surrounding communities. Clearly, colleges by themselves cannot optimally reduce the alcohol problems among college students or other persons the same age not in college, and binge drinking and driving under the influence of alcohol has also increased among this group. For every 18- to 24-year-old college student, there are two persons their age not in college. Entire community norms and policies need to be changed. If colleges adopt stricter alcohol policies and enforcement but surrounding communities do not, the college student alcohol problems will simply be pushed off campus into surrounding communities. Conversely, if communities tighten their policies and enforcement and colleges do not, it will only push the problems back onto college campuses. Also, many of the problems experienced as a result of excessive student alcohol consumption can affect people other than the college drinkers themselves.

Further, many college students develop problematic drinking habits before they enter college. Analyses of the National College Alcohol Survey indicate the younger college students were when they first drank to intoxication, the greater the likelihood that they experience alcohol dependence, rode with drinking drivers, drove after drinking, were injured under the influence of alcohol, and had unplanned and unprotected sex after drinking. The recently published analyses of elementary and middle school interventions that produced reductions in drinking and related problems that carried over to young adult years suggests these approaches might help further reduce in college drinking and related problems. This warrants research attention.

Consequently, colleges and surrounding communities need to work together to implement multi-faceted programs at various levels. They need to expand the reach of individually oriented counseling programs to include university health services and web-based outreach, test and refine promising new approaches to involve parents, and test and further implement policies that can further reduce college age and underage alcohol misuse.

Collectively, they need to involve multiple departments of city government as well as concerned private citizens and organizations and multiple sectors of the college community, presidents, deans, other administrators, campus security, residence counselors, health service providers, alumni, faculty, and students if they want to most effectively reduce harmful drinking and the myriad of health and social problems linked to harmful drinking.

SEE ALSO

Alcohol Use Disorders, Binge Drinking, Overdose, Adolescent Substance Use: Symptoms and Course, Epidemiology of Adolescent and Young Adult Alcohol,

Tobacco, and Drug Use and Misuse in the United States, Alcohol and Drug Use in Sexual Minority Youth and Young Adults (lesbian, gay, bisexual, transgender), Athletes and Substance Use, Drinking Patterns, Alcohol Consumption, and Aggressive Behavior, Alcohol and Sexual Violence, Alcohol, Sexual Risk Taking, and Sexually Transmitted Infections

List of Abbreviations

BAC blood alcohol concentration
BMIs brief motivational interventions
DUI driving under the influence
DWI driving-while-intoxicated
ISFP Iowa Strengthening Families Program
NESARC National Epidemiologic Survey on Alcohol and Related Conditions
NIAAA National Institute on Alcohol Abuse and Alcoholism
NLAES National Longitudinal Epidemiologic Survey
PDFY preparing for the drug free years
SAMHSA Substance Abuse and Mental Health Services Administration

Further Reading

Carey, K.B., Scott-Sheldon, L.A.J., Carey, M.P., et al., 2007. Individual-level interventions to reduce college student drinking: a meta-analytic review. Addictive Behaviors 32, 2469–2494.

DeJong, W., Larimer, M.E., Wood, M.D., 2009. College drinking: new research from the National Institute on Alcohol Abuse and Alcoholism's Rapid Response to College Drinking Problems Initiative. Journal of Studies on Alcohol and Drugs (Suppl. 16), 5–147.

Elder, R.W., Lawrence, B., Ferguson, A., et al., 2010. The effectiveness of tax policy interventions for reducing excessive alcohol consumption and related harms. American Journal of Preventive Medicine 38 (2), 217–229.

Fell, J., Fisher, D.A., Voas, R.B., et al., 2009. The impact of underage drinking laws on alcohol-related fatal crashes of young drivers. Alcoholism: Clinical and Experimental Research 33 (7), 1–12.

Fleming, M.F., Balousek, S.L., Grossberg, P.M., et al., 2010. Brief physician advice for heavy drinking college students: a randomized controlled trial in college health clinics. Journal of Studies on Alcohol and Drugs 71 (1), 23–31.

Harris, S.K., Sherritt, L., Van Hook, S., Wechsler, H., Knight, J.R., 2010. Alcohol policy enforcement and changes in student drinking rates in a statewide public college system: a follow-up study. Substance Abuse Treatment, Prevention, and Policy 5, 1–11.

Hingson, R., Heeren, T., Zakocs, R., et al., 2003. Age of first intoxication, heavy drinking, driving after drinking and risk of unintentional injury among U.S. college students. Journal of Studies on Alcohol 64 (1), 23–31.

Ichiyama, M.A., Fairlie, A.M., Wood, M.D., et al., 2009. A randomized trial of a parent-based intervention on drinking behavior among incoming college freshmen. Journal of Studies on Alcohol and Drugs (Suppl. 16), 67–76.

Larimer, M.E., Cronce, J.M., 2007. Identification, prevention, and treatment revisited: individual-focused college drinking prevention strategies 1999–2006. Addictive Behaviors 32, 2439–2468.

Moreira, M.T., Smith, L.A., Foxcroft, D., 2009. Social norms interventions to reduce alcohol misuse in university or college students (review). Cochrane Database of Systematic Reviews 3 CD006748.

National Institute on Alcohol Abuse and Alcoholism (NIAAA), 2002. Task Force of the National Advisory Council on Alcohol Abuse and Alcoholism. A Call to Action: Changing the Culture of Drinking at U.S. Colleges (NIH publication no. 02–5010). National Institutes of Health, U.S. Department of Health and Human Services (DHHS), Bethesda, MD.

Nelson, T.F., Toomey, T.L., Lenk, K.M., Erickson, D.J., Winters, K.C., 2010. Implementation of NIAAA College Drinking Task Force recommendations: how are colleges doing 6 years later? Alcoholism: Clinical and Experimental Research 34 (10), 1–7.

Saltz, R.F., Paschall, M.J., McGaffigan, R.M., Nygaard, P. Alcohol risk management in college settings: the Safer California Universities Randomized Trial. American Journal of Preventive Medicine, in press.

Shults, R.A., Elder, R.W., Sleet, D.A., et al., 2001. Reviews of evidence regarding interventions to reduce alcohol-impaired driving. American Journal of Preventive Medicine 21 (4), 66–88.

Spoth, R., Trudeau, L., Guyll, M., Shin, C., Redmond, C., 2009. Universal intervention effects on substance use among young adults mediated by delayed adolescent substance initiation. Journal of Consulting and Clinical Psychology 77 (4), 620–632.

Wechsler, H., Nelson, T.F., 2010. Will increasing alcohol availability by lowering the minimum legal drinking age decrease drinking and related consequences among youths? American Journal of Public Health.

CHAPTER

63

Drinking Patterns, Alcohol Consumption, and Aggressive Behavior

Brian M. Quigley, Kenneth E. Leonard

Research Institute on Addictions, Buffalo, NY, USA

OUTLINE

ALCOHOL AND AGGRESSION

The co-occurrence of alcohol use and physically aggressive behavior has been observed throughout recorded history and across numerous cultural contexts. In the 1400s, Netzahualcoyotl of Texcoco described alcohol as a cause of violence among family members. In the mid-1700s, William Hogarth issued two prints, one of which, Gin Lane, depicted the pernicious effects of gin, including two men attacking each other in front of a distillery. Throughout the 1800s, alcohol was described in morality tales of the temperance movement as well as in specific court cases as leading men and women to violence. In one of case, a man in Middletown, Connecticut, was hung for beating his wife to death after 2 days during which both he and his wife had been drinking, arguing, and fighting with each other. Despite these destructive effects of alcohol, there has also been a recognition that most instances of alcohol use are accompanied by positive effects. The challenge for research has been to understand whether alcohol leads to aggressive behavior, and if so, for whom and under what circumstances. The co-occurrence of alcohol use and physically aggressive behavior has been examined in numerous domains including crime statistics, survey research of general and domestic violence, and experimental studies.

DEFINITIONS AND DISTINCTIONS

In this chapter, "aggression" is defined as any behavior with the intent to cause harm to another person. This term often includes behaviors with the intent to cause psychological harm, such as psychological, verbal, or emotional aggression, but in general, the term "aggression" refers to the intent to cause physical harm. The term "violence" usually refers to more severe acts of physical aggression that have the potential for injury or death, as well as acts that have been judged as "criminal."

Several distinctions regarding alcohol also need to be made. The intoxicating effects of alcoholic drinks should be considered acute effects and distinguished from

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00063-2

chronic effects, which refer to effects that occur only after the individual has used the substance for some time. Chronic effects associated with long-term alcohol use include, but are not limited to, physical damage, such as liver cirrhosis and changes in executive cognitive functioning, and social damage, such as disrupted friendships and marriages. Similarly, when discussing the association between alcohol use and aggression, we need to distinguish substance user from substance use. Although alcohol use may show certain relationships with aggression, possibly due to the acute effects of alcohol, an individual who is a substance user may be more likely to engage in aggressive behavior for various other reasons, some of which may have little to do with the alcohol itself. An individual who is a heavy drinker may differ in a variety of important ways from an individual who is not. These may include individual differences in impulsiveness and cognitive functioning as well as differences in the typical social and physical environment in which drinkers or heavy drinkers find themselves.

ALCOHOL USERS AND VIOLENCE

Studies of violent criminals have often found that they are very likely to be heavy drinkers, to have experienced alcohol problems, or to have an alcohol disorder. Based on data from the National Survey of Jail and Prison Inmates, 50% of homicide offenders were classified as "heavier" drinkers (2–5 drinks per day) or "much heavier" drinkers (more than 5 drinks per day). The use of offender populations (i.e. individuals arrested and convicted for the commission of a violent crime) has the strength of ease and efficiency. It is very convenient inasmuch as the individuals have been already been identified and located. In addition, the violent behavior of interest has been documented by evidence and by the judicial process, and it is not reliant on the individual admitting responsibility. Studies of incarcerated offenders have provided fairly clear results regarding the prevalence of alcohol and substance use/abuse among violent and nonviolent offenders. In general, studies of incarcerated criminals have found that the typical drinking patterns of violent and property offenders were very similar, although this is not true for the use of other drugs.

Although offender populations provide a useful and convenient sample from which to collect data, these populations also have limitations. Offender populations reflect a minority of individuals who have behaved violently. Many crimes of violence go unreported or unsolved. In comparing police reports to victims' reports, it has been found that only 4% of the violent incidents reported by victims come to the attention of the police. Consequently, the characteristics of violent offenders may reflect both the causes of violence as well as the factors that expedite arrest and conviction. For example, the common observation that about 50% of homicide offenders were drinking at the time of the homicide may reflect the inability of intoxicated offenders to avoid detection. Second, the selection of nonviolent individuals from incarcerated populations can be difficult. Although one can rely on the crime of conviction to identify the violent offenders (i.e. homicide offenders, assault offenders), conviction of a nonviolent crime does not ensure that the sample is not violent. The nonviolent conviction may be the result of a plea bargain for a violent offense, or it may reflect the most serious charge that the district attorney could substantiate. Also, many individuals engage in both violent and nonviolent crimes. Thus, many of the individuals convicted of a nonviolent crime have committed violent crimes in the past. Despite a variety of methodological problems with much of this research, it is clear that violent criminals are very likely to use alcohol excessively. The critical issue to recall is that data indicating substance users and/or abusers are likely to engage in violence do not necessarily suggest that the substance use or even substance abuse caused the violence.

Despite the difficulty and cost, studying the occurrence of alcohol and violence in the general population is an important approach to identifying a sample of violent individuals. For example, The Epidemiology Catchment Area (ECA) project studied the prevalence of psychiatric disorders in the general population as well as information regarding involvement in violent episodes. Evidence from analysis of these data demonstrated that 15.3% of individuals with alcohol problems had been involved in violent incidents. Only 2% of the individuals without alcohol abuse problems were involved in violent incidents. Another general population survey, the Household Survey of Drug and Alcohol Abuse found similar results. Among those with no alcohol use in the previous year, only 2.7% were involved in violence. When the respondents drank at least monthly, 14.6% had experienced violence in the previous year. Meta-analyses of more than 100 case–control and cross-sectional studies in the general population have reported significant associations between drinking patterns and criminal violence and between drinking patterns and intimate partner violence.

Studies of general populations have also been conducted in several other countries. The association has been found in large general population studies in Great Britain, Canada, Australia, and the United States. Multicountry studies conducted by the World Health Organization found associations in population surveys conducted in Central and South America (Brazil, Columbia, Venezuela, Costa Rica, and El Salvador),

Asia (India and Indonesia), Europe (Spain), and Africa (South Africa). Other multinational studies and smaller studies have expanded the countries in which the relationship has been observed to China, Cambodia, Columbia, Dominican Republic, Haiti, Nicaragua, Peru, Bangladesh, Uganda, Ethiopia, Egypt, Nigeria, and Iran. Thus, the relationship between heavy drinking and violence has been observed in many diverse cultures.

Research has also been conducted among more select samples. Studies have observed the relationship between drinking and violence among alcoholics, battered women, and men arrested for partner violence. Studies in health clinics and emergency rooms, which tend to identify victims of violence, have also consistently found that heavy drinking was associated with the occurrence or severity of violence. Recent research in emergency rooms in Argentina, Australia, Canada, Mexico, Spain, and the United States found that individuals injured in a violent event showed heavier drinking patterns than those injured in an accident.

Although much of the research has used cross-sectional designs, a small number of studies have examined the relationship between alcohol and violence using longitudinal designs. These longitudinal studies have mainly focused on aggressive behaviors in the context of intimate partners. Several studies have examined the relationship over periods of 1 to 2 years. These studies have almost consistently found that men's drinking was longitudinally predictive of men's aggressive behavior toward their spouse, either as a main effect or in certain subgroups of men (e.g. hostile). However, none of these longitudinal studies have found that women's drinking was related to their own aggressive behavior nor to their husband's aggressive behavior. In addition, a number of treatment studies have found that alcoholics who maintain sobriety after treatment engage in less partner violence in the year after treatment compared with alcoholics who relapse. Studies that have examined alcohol and partner aggression over longer periods of time have been less consistent, as have the few longitudinal studies of alcohol and general (not partner) aggression.

In addition to the general finding of an association between drinking patterns and violence, several additional findings are important to note. First, across the cross-sectional and longitudinal studies, a wide variety of measures of drinking patterns have been used, including average daily use, frequency of heavy drinking, alcohol problems, and alcohol use disorders. Some research has found that measures of current alcohol problems are related to current violence but measures of lifetime problems are not. Moreover, research suggests that indicators of heavy episodic drinking are most consistently associated with aggression but measures of the typical frequency of drinking are often not associated with aggression. Second, although the relationship between drinking patterns and aggression has been observed among men and women, the results are less consistent with respect to women. Most of the research on women's aggression has focused on intimate partner violence. Studies in general population samples have often failed to find an association between drinking patterns and women's violence toward a partner, or have found that a significant univariate relationship is no longer significant after controlling for the male partner's drinking pattern. In contrast, an association between women's drinking pattern and aggression toward her partner often emerges in studies of clinical samples, perhaps suggesting that some of the characteristics of women in clinical samples may potentiate the association between drinking patterns and violence.

ALCOHOL USE AND AGGRESSIVE EVENTS

The earliest and most well-known approach to understanding the relationship between alcohol and aggression was to examine criminal events. Since the 1950s, crime statistics have consistently shown a co-occurrence of alcohol use and criminal behavior. It has been estimated that approximately 60% of violent crimes involve alcohol use, with estimates ranging from 24 to 82% of perpetrators in assaults having consumed alcohol prior to committing the crime. According to data from the National Survey of Jail and Prison Inmates, 36% of the offenders had consumed only alcohol prior to committing a homicide and an additional 13% had consumed alcohol in combination with some other drug. As only 7% of the offenders had consumed drugs but no alcohol prior to the homicide, these data show that alcohol use commonly precedes the occurrence of violent actions, but the use of other substances is less common. However, the use of alcohol prior to the crime of conviction was more common for violent offenses than property offenses.

Although these statistics seem impressive, there are a number of caveats regarding their interpretation. Although we can estimate the frequency of alcohol use in episodes of violence, we do not know the frequency of alcohol use in nonviolent interactions. Finding that 50% of assaults were accompanied by alcohol consumption by the offender would not be of interest if 50% of the time that two individuals got together, one or the other had been drinking. Even though we do not have information concerning the presence of drinking in nonviolent get-togethers, it seems unlikely that it would approach 50%. However, it is also clear that not

everyone is equally likely to become involved in an assault. We know that violence is more common among young adults than among older adults, and among men than among women. There are also individual difference factors that place some individuals at higher risk for assault than others, such as antisocial personality characteristics. Perhaps 50% of the time that antisocial young men get together, one or the other has been drinking. If this were the case, the fact that 50% of assaults occur under the influence of alcohol would not be impressive. The point of this discussion is that simple rates of involvement are interesting, but do not provide meaningful information concerning associations of substance use and violence. Studying the involvement of alcohol use in aggressive events requires research designs that can assess the presence of alcohol in events differing in aggression, such as no aggression versus aggression, or moderate aggression versus severe aggression. Alternatively, designs that can examine the occurrence of aggression in events varying in alcohol use can also provide information about the association. In either of these approaches, it is of critical importance to control for individual characteristics that might be associated with the type of event under investigation.

One approach has been to compare aggressive events involving some individuals with control events, such as verbal conflict, involving other individuals. Because people who have never experienced an aggressive event are likely to differ from those who have, individual differences need to be controlled for statistically. Although the earliest research using such a "between-subjects" approach did not find a relationship between alcohol and assaults, more recent research has. In one recent study, observers of police–citizen encounters found that alcohol consumption was more common in instances of spousal assault and other violent crimes than in instances of nonviolent police–citizen encounters. In a study of intimate partner aggression among newlywed couples, alcohol consumption by the husband was more common among episodes of moderate and severe aggression than among episodes of verbal conflict. Alcohol use has also been found to be more common in more severe episodes of violence in contrast to less severe episodes, even when these contrasts involved homicide versus nonlethal violence.

A second approach to examining alcohol and aggression in events has been to compare aggressive events and less aggressive events reported by the same person, the "within-subject" approach. This could be accomplished by comparing alcohol use before an event among people who reported both a verbal and physical aggressive event. One might examine all the individuals who report both a physical violence and a verbal aggression event, and compare the alcohol consumption preceding these two events. In studies of young married couples, drinking by the husband has been found to be more common in a violent event than in an argument. In one study, youthful offenders reported on an assault that they committed that resulted in an injury and the last time that they were with the same people at the same time and on the same day of the week. The assaultive event and the control event were very similar in terms of the context of the event, but the offenders reported drinking more prior to the assaultive event than prior to the control event. In a similar vein, in a study of conflict events occurring in bars, alcohol consumption did not predict the occurrence of aggression, but it was associated with aggression severity and physical harm, and this relationship was present after controlling for personality and situational factors. In addition, in treatment samples, the level of alcohol consumption has been found to be higher prior to violent events in contrast to verbal aggression events.

In recent years, rather than collecting data on two discrete episodes differing in the level of violence (or level of alcohol consumption), researchers have begun to collect detailed data regarding drinking, conflict, and aggression every day over the course of 2–3 months. In some studies, this has been collected retrospectively using the Time Line Followback procedure, an interview designed to collect data on drinking as well as other events on a daily basis over an extended period of time. Other studies have used similar interviews focusing on the preceding week. The most sophisticated approach has involved daily reports regarding drinking and the occurrence of aggression in the preceding 24 h. All of these approaches allow for the use of highly sophisticated multilevel models to examine the relationship between alcohol and violent events. This approach allows the researcher to examine whether drinking on one day was associated with the occurrence of aggression the same or the following day. This approach has found associations at the daily level and the occurrence of aggression in a number of diverse samples. One study conducted weekly interviews of mentally ill individuals at high risk for violence and found that alcohol use was predictive of violence the same day and the next day, but that violence was not predictive of subsequent alcohol use. Among college students reporting on drinking and a variety of different behaviors, the students estimated blood alcohol concentration (BAC) on a specific day was associated with the occurrence of aggression and vandalism on the same day. In a recent study, drinking on one day was linked to aggression against members of a sexual minority group, controlling for drinking patterns and other risk factors for aggression. In short, across a variety of event-based methodologies, heavier alcohol use on a given day is clearly associated with aggressive behavior on that specific day, and sometimes is predictive of aggression on the subsequent day.

EXPERIMENTAL STUDIES

Some of the best evidence for the causal effect of alcohol on aggressive behavior comes from laboratory studies. In such studies, participants may ostensibly be allowed to administer an aversive stimulus to an individual who is believed to be another subject but is, in actuality, an experimental assistant or confederate. The aversive stimulus may be physical (such as noise or electric shock), verbal (such as an evaluation), or symbolic (as in a response that reduces the money that can be earned in a task). These behaviors are viewed as an analog for real-world aggressive behavior and although they are an underestimate of the actual behavior of interest, they are generally accepted as reasonably valid measures of aggressive responding. Laboratory studies are important because they allow for the precise examination of acute alcohol effects on aggressive behavior. The pharmacological effects of substances are often dose related, and tied to the amount of the substance in the brain, which is indexed by their breath or blood alcohol concentration. In the natural environment, individuals differ with respect to the amount of alcohol they drink, the alcohol content of the beverage, the duration and speed of their drinking, and the activities that they engage in while drinking. Each of these can have an impact on the person's actual BAC. In the laboratory, these factors can be controlled and a more precise BAC can be achieved by participants. Laboratory studies allow for control over extraneous variables and provide the ability to rule out alternative explanations for findings allowing the conclusion that one variable has caused the changes observed in another variable. In the case of alcohol, through the inclusion of a placebo condition, the use of laboratory studies permits the assessment of any effects that occur simply on the basis of the individual's expectations concerning the drug. Experimental paradigms allow for the inclusion of situational variables that allows us to specify the conditions under which alcohol does and does not result in increased aggression. These factors, coupled with the ability to evaluate individual difference factors involved in the relationship, allow for the testing of theoretical explanations of the relationship.

The most common procedure used in the laboratory when dealing with alcohol and aggression has been the Taylor reaction time procedure (TAP). In this procedure, participants play a reaction time game with another participant (usually a confederate of the experimenter). For each trial, the subject sets a level of shock that they wish to deliver to the other person. The loser of the reaction time trial (i.e. the slower reactor) then gets shocked. The game is rigged so that the participant receives some shocks (usually in an escalating manner) and is allowed to deliver shocks on others on other trials. The first trial is sometimes considered a measure of proactive or instrumental aggression and subsequent trials a measure of reactive aggression.

Meta-analyses of the laboratory literature have concluded that, in general, the administration of alcohol does lead to increased levels of aggression. Higher levels of alcohol intoxication (BAC = 0.08 to 0.10) have been found to lead to more aggressive responding in the TAP than low levels of intoxication (with effect sizes as estimated by Cohen's d ranging from 0.47 to 0.61). In one review of 30 laboratory studies investigating the effects of alcohol on aggressive responding, there was significantly more aggression when comparing alcohol administration conditions with no alcohol control conditions as well as to placebo conditions, suggesting that alcohol has a psychopharmacological effect beyond any effect of believing that one has consumed alcohol. Several studies have found that high doses of alcohol result in higher levels of aggression than low doses of alcohol. In addition, the relationship between alcohol dose, at least at the levels given in the laboratory, was found to be linear in a study in which various alcohol doses were given to participants (0.0 g kg^{-1}, 0.125 g kg^{-1}, 0.25 g kg^{-1}, 0.50 g kg^{-1}, 0.75 g kg^{-1}, and 1.00 g kg^{-1}).

These meta-analyses of laboratory studies have also typically failed to find that placebo conditions increase aggression over no alcohol control conditions. Although placebo effects in alcohol and aggression research are rare, recent evidence suggests that alcohol-related cues can instigate aggressive responding even in the absence of alcohol consumption. In two recent studies, when individuals were exposed to alcohol-related cues, even subliminally, they tended to respond faster to aggressive words and to be more aggressive toward a person who had insulted them. This effect was also largest among those holding the expectancy that alcohol makes them aggressive. These findings, in contrast to the negative placebo findings, suggest that alcohol expectancies may have effects on aggression although other factors may ameliorate the effects when actually consuming alcohol. It is one of the current contradictions in the data that will require additional research to untangle.

In the TAP procedure, it is possible to examine the effect of alcohol on aggression in relation to the provocations of the other person (confederate) in the experiment. This may be accomplished through the verbal behavior of the participant or through feedback to the participant that the other person is increasing the level of aversive stimuli to be administered to the participant. Studies using verbal behavior suggest that alcohol does not increase aggressive behavior relative to no alcohol controls when there are verbal behaviors that reduce the aggressive threat of the other person, such as a declaration by the other person that he intends to select only low shock levels. Similarly, alcohol does not seem to result in

increased aggression in the presence of other situational cues that press for nonaggressive behavior, such as a nonaggressive norm or pressure from another person in the room to behave nonaggressively, even in the face of increased attack from the other person. In contrast, alcohol does lead to increased aggression in the presence of aggressive instigations. For example, participants who have received alcohol increase their aggression against a nonaggressive opponent if an observer pressures them to do so, but sober participants do not. Similarly, intoxicated participants increase their use of an extreme level of shock more than sober participants in response to an attempt by their opponent to administer that extreme shock to them. However, when the provocative cues are at their highest level, the difference between alcohol and control conditions is often attenuated. For example, after overhearing their opponent indicate an intention to use only the highest shock settings, intoxicated and sober participants are equally aggressive. Similarly, in many studies, the opponent gradually increases the level of shocks selected for the participant. As this provocation increases, the difference between the alcohol condition and the control condition decreases.

Laboratory research has also investigated the individual differences that may moderate the effect of alcohol on aggressive behavior. Among the individuals, the differences that have been found to moderate the effects of alcohol on aggressive responding are empathy, irritability, dispositional aggressiveness, trait anger, self-control, and cognitive executive functioning. In general, it appears that alcohol has stronger effects among individuals with higher levels of traits that facilitate aggression and among individuals with lower levels of inhibitory characteristics. For example, alcohol facilitates aggression among those who are high in anger and low in empathy. It should be recognized that most of this research focuses on one specific moderator at a time, and virtually none of the research has examined individual differences and situational characteristics in the same study. In addition, although research very rarely finds placebo effects (which are meant to control for expectancies regarding alcohol), individual differences in expectancies regarding alcohol's effects on aggression have sometimes been found to be related to aggressive responding. In one study using the TAP, individual differences in alcohol expectancies regarding alcohol interacted with alcohol consumption such that those with high expectancy that alcohol makes them aggressive, who consumed alcohol and who were provoked, were more likely to respond with the most extreme reactions. Recent research has indicated that the belief that alcohol makes one aggressive does lead to greater aggressive responding when given alcohol in the laboratory, although this effect seems to disappear when controlling for trait aggressiveness.

EXPLANATIONS FOR THE RELATIONSHIP BETWEEN ALCOHOL AND AGGRESSION

Noncausal Explanations

One set of explanations for the alcohol/aggression relationship suggest that the use of alcohol does not precipitate violence, rather the use of alcohol and the commission of violent acts are precipitated by a common set of circumstances. The relationship between alcohol use and violent behavior only exists because some third variable is the actual cause of both behaviors. According to the general deviance syndrome, individuals who are likely to engage in violent behavior also demonstrate predispositions to engage in other deviant behavior. These subcultures are most likely to exist among lower socioeconomic classes, especially among men, and are distinguished by a concern for a masculine identity and engaging in exciting or dangerous activities. This explanation can be seen in the findings on violence in bars indicating that certain types of aggressive people combined with certain types of permissive bars are the recipe for violence. College-age males view the bar environment as a place of sexual competition and rivalry as well as a place to display one's masculinity. It has been found that those with a history of substance dependence are more aggressive in laboratory measures of aggression than those who have no history of substance dependence. In addition, early onset of aggressive behavior in boys has been found to be predictive of later violence, criminal behavior, and alcohol use. In addition individuals with certain personality characteristics may be more likely to engage in aggression and use alcohol, for example those high on sensation seeking or impulsivity, or low on self-control or other aspects of executive cognitive functioning.

Alcohol may also be related to aggression because of the context of typical alcohol consumption. One of the most frequent locations in which individuals report experiencing violence is in or around bars. In one recent survey, over 40% of males and females who reported that they had experienced physical violence in the past year reported that the most severe incident had occurred in or around a bar. In a separate survey of young adults in the Albany, NY, area, young males who engaged in more nightlife activities were more likely to be targets of violence and to have experienced non-family-related violence. Nightlife activities are often permeated by a culture of masculinity. Drinking itself is used to promote one's masculinity with "real men" being able to hold their liquor and being able to have one up on others in the pub. An analysis of drinking occasions in Sydney, Australia, demonstrated that physical aggression while drinking was motivated by personal pleasure

and by a desire to present or create a masculine identity. In addition, much violence is also motivated by the carnival atmosphere at some drinking establishments. It both provides a show for the clientele and a pleasurable diversion for those who wish to engage in the behavior. Thus, the clientele, the bar environment, and the effects of alcohol all play roles in generating bar violence. When comparing bars in which violence was common with bars in which violence was infrequent, there were several differences related to both the clientele and to the bars themselves. Clientele at violent bars were more likely to report alcohol problems and more likely to score higher on individual differences in anger. Bars at which violence was more common were noisier, warmer, employed bouncers, disregarded illegal activity, and generally provided a more permissive atmosphere. The differences in clientele were also predictive of the bar characteristics, suggesting that certain types of people choose certain types of bars, which leads to violence in some places but not in others.

Those committing crimes may use alcohol and drugs in order to help them prepare themselves for the act. Ritual behaviors engaged in by street gangs, such as the induction of a new member or the expulsion of old members, often involved the use of alcohol and drugs as well as the commission of violent or criminal acts. Alcohol and drugs are used by Chicano gang members in Los Angeles to relax before engaging in acts of violence. To prepare themselves for conflict with other gangs, the members drank alcohol and smoked phencyclidine-laced cigarettes. Associating alcohol and drug use with going into a dangerous situation is not restricted to street gangs. Male college students generally associate heavy alcohol consumption with masculinity. According to some research, up to two-thirds view alcohol as a fuel for strength and aggression and believe that alcohol induces courage or invincibility in men.

It is clear that alcohol and drugs may be indirectly associated with the occurrence of aggression. However, the alcohol/aggression relationship cannot be totally explained by these types of mechanisms. They do not explain the laboratory research in which these variables are controlled and yet alcohol administration leads to higher levels of aggression. In order to more fully explain the relationship between alcohol and aggression, we must examine other mechanisms.

Expectancy Explanations

In general, alcohol expectancy explanations for the relationship between alcohol and aggression take three forms. The first is the deviance disavowal hypothesis. According to this perspective, periods of intoxication may be socially defined as a "time out" from usual social rules in which one can engage in behavior not normally sanctioned. The preponderance of the evidence suggests that alcohol use, at times, may lead one to be seen as less blameworthy, but overall it does not support the use of alcohol as an excuse for aggression. More often, alcohol use has the opposite effect and it increases attributions of blame to the drinking perpetrator. When evaluating the deviance disavowal hypothesis, it is important to recognize that attributions regarding blame are not simple and that one is not simply engaging in attributions for the behavior enacted while intoxicated, but for the act of intoxication itself. Although alcohol may have played a causal role in a person enacting the negative behavior, the person may be seen to be blameworthy for becoming intoxicated and ultimately to blame for the behavior while intoxicated. Consequently, it does not seem that individuals generally use intoxication as an excuse for engaging in antinormative behavior such as aggression.

The second expectancy explanation for the relationship between alcohol and aggression suggests that alcohol expectancies provide a guide or script as to how to behave when drinking. The belief that alcohol makes one aggressive has been found, somewhat inconsistently, to be correlated with measures of aggressive behavior. Individual differences in this expectancy have been found to be related to trait measures of aggression and to self-reports of aggressive behavior. They have also been predictive of reports of intimate partner violence. In laboratory studies, individual differences in alcohol expectancies have inconsistently been found to lead to more aggressive responding and alcohol placebos, which are meant to manipulate alcohol expectancies, rarely have an effect on aggressive responding (as cited above).

Currently, a newer interpretation of the alcohol expectancy concept has provided a third explanation based on a more cognitive model of expectancies combined with the General Aggression Model. Alcohol expectancies, rather than just being considered beliefs or attitudes, can also be conceptualized as a cognitive associative network in memory involving both explicit and implicit cognition and affect. Explicit measures such as self-report measures may underestimate the construct and, in addition, are susceptible to social desirability biases. In this model, the activation of certain expectancies subsequently activates other concepts, memories and scripts. Alcohol concepts are closely associated with aggression in some people but not in others. The expectancies activated in any given situation are dependent on cues in the environment as well as on the person's cultural background and learning history. For those with a close association between alcohol and aggression, an alcohol cue may be all that is needed to give rise to aggressive thoughts and scripts. The aggressive cue would likely also serve to orient the person toward the aggressive aspects of the situation, and once a cognitive

associative network related to aggression has been activated, it influences the processing of subsequent stimuli. Current research showing that subliminal alcohol-related cues give rise to the activation of aggressive thoughts and behaviors in individuals holding the expectancy that alcohol causes aggression fits with this model.

Pharmacological Explanations

Because alcohol can have many effects including making one relaxed, happy, or sleepy, there is no reason to assume it will automatically make one aggressive. As a consequence, the pharmacological explanations for alcohol's relationship with aggression have to do with how alcohol influences cognitive processes. The pharmacological explanations for alcohol's effects on aggression emphasize alcohol's influence on what cues are attended to in the environment. For example, when a mirror is placed in front of intoxicated individuals, they attend to self-relevant cues more than if the mirror had not been there. The question of why some cues are attended to and others are not has been discussed by numerous theories including alcohol myopia, anxiolytic effects of alcohol, and the social information processing theory.

The alcohol myopia theory posits that alcohol interferes with the processing of attentional information, causing one to attend to the most salient aspect of the situation. When both instigating and inhibiting responses are equally likely, intoxication causes an individual to attend to instigating cues more so than to inhibitory cues, a situation known as "inhibition conflict." In a meta-analysis of several different types of social behaviors including aggression, risk-taking, self-disclosure, and sexual interest, it was demonstrated that alcohol's effects on all of these social behaviors only occurred when situations of inhibition conflict existed. However, none of the studies explicitly examines manipulated inhibition conflict. The authors used a post hoc categorization of the studies to identify situations of inhibition conflict. Some primary research supportive of the myopia hypothesis has been reported. Additional research has shown that alcohol myopia effects predict numerous types of behavior including prosocial behavior, positive self-evaluations, as well as decisions about drinking and driving and condom use. More current studies using the TAP to explicitly study intoxicated aggression have shown that when intoxicated participants are distracted by another cognitive task, the relationship between intoxication and aggression is reduced, suggesting that when attention is diverted from aggression-facilitating cues, intoxicated individuals are less likely to respond aggressively.

However, alcohol myopia is not the only theory as to why alcohol makes one attend to facilitative cues and disregard inhibitory cues. The anxiolysis–disinhibition model proposes that in situations in which aggression has the potential to occur, anxiety is also present. This anxiety can be due to fear of retaliation for engaging in aggressive actions or fear of censure for aggressive actions. The anxiety arising from the situation may inhibit responses to it. Hence, in situations in which aggression is likely and alcohol is also present, those who have been drinking alcohol feel less anxiety than those who have not been drinking, and are therefore more likely to engage in aggressive behavior. A meta-analysis examining the myopia hypothesis versus the anxiolytic hypothesis found evidence supporting both mechanisms. In terms of alcohol myopia, a significant effect of alcohol on aggressive behavior was found in situations of high inhibition conflict (as rated by the researchers). A small but significant effect was also found in situations of low inhibition conflict. However, the effect of alcohol was stronger for high versus low inhibition conflict suggesting a significantly stronger effect under situations of high inhibition conflict. In addition, the influence of inhibition conflict depended on the dose of alcohol. When low dosages were used, a significant effect of alcohol on aggression was found in both low and high inhibition conditions. However, when high doses of alcohol were used, the effect of alcohol was present only under high conflict inhibition conditions. The low inhibition condition did not produce a significant effect. In support of the anxiolysis–disinhibition model, the difference between intoxicated and sober individuals in aggressive behavior increased as the amount of anxiety present in the studies increased. Although these results suggest that inhibition conflict and anxiolysis–disinhibition do play roles in the effect of alcohol on aggressive behavior, it is unclear if they are mutually exclusive hypotheses or really part of the same mechanism.

One other model has been used to explain the effects of alcohol on aggressive behavior. The social information processing perspective argues that aggressive actions are a product of several separate cognitive skills, including the encoding and interpretation of cues existing in the interpersonal situation, response generation, outcome evaluation, response selection, and response enactment. In one study, participants who consumed or did not consume alcohol were presented with a video tape showing one individual rudely changing the television channel on another person. Different versions of the video showed the individual behaving aggressively or in a more neutral manner. Participants were then asked to rate the individuals on several dimensions and then decide what would happen next. Results showed that alcohol did not have an effect on participants' perceptions of the situation. Intoxicated individuals did not view the offending individuals' action as any more intentional or aggressive than sober participants.

However, there was a difference between intoxicated and sober subjects in terms of decision making. Those who had consumed alcohol made less adaptive (i.e. more aggressive) response options and were more likely to endorse the use of aggressive options then were subjects who received a placebo.

The evidence that alcohol does have pharmacological effects on aggressive behavior is substantial. It is clear that alcohol influences the cognitive decision-making ability of individuals; however, the mechanism for that effect is still not well understood. Alcohol myopia, anxiolysis–disinhibition, and social information models all have some support in the literature. It is not yet clear which is the best explanation. It is possible all three may actually be aspects of a single mechanism. As mentioned earlier, alcohol myopia and anxiolysis–disinhibition are similar in the mechanisms they propose and in many hypotheses. The social information model is not necessarily at odds with either the myopia or the anxiolysis–disinhibition models. In addition, the current evidence for some people who hold expectancies about alcohol, the presence of alcohol-related cues in the absence of intoxication giving rise to aggression suggests that a more cognitive perspective integrating all these models may be the most complete explanation for the acute effects of alcohol on aggressive behavior.

SEE ALSO

Families and Addiction, Substance Induced Myopia

List of Abbreviations

BAC blood alcohol concentration
IPV intimate partner violence
TAP Taylor aggression procedure

Glossary

Aggression any behavior performed with the intent to harm another person.
Alcohol myopia the state of tunnel vision or the limiting of attention caused by intoxication.
Anxiolysis-disinhibition model the process by which inhibitions are reduced because intoxication reduces the anxiety associated with said inhibitions.
Anxiolytic a substance that pharmacologically reduces anxiety.
BAC blood alcohol concentration.
Deviance disavowal the concept that alcohol consumption provides a time out from normal social rules and allows one to avoid blame or punishment for antinormative behaviors.

Further Reading

Bushman, B.J., Cooper, H.M., 1990. Effects of alcohol on human aggression: an integrative research review. Psychological Bulletin 107, 341–354.
Duke, A.A., Giancola, P.R., Morris, D.H., Holt, J.C.D., Gunn, R.L., 2011. Alcohol dose and aggression: another reason why drinking more is a bad idea. Journal of Studies on Alcohol and Drugs 72, 34–43.
Foran, H.M., O'Leary, D.K., 2008. Alcohol and intimate partner violence: a meta-analytic review. Clinical Psychology Review 28, 1222–1234.
Giancola, P.R., 2007. The underlying role of aggressivity in the relation between executive functioning and alcohol consumption. Addictive Behaviors 32, 765–783.
Giancola, P.R., Josephs, R.A., Parrott, D.J., Duke, A.A., 2010. Alcohol myopia revisited: clarifying aggression and other acts of disinhibition through a distorted lens. Perspectives on Psychological Science 5, 265–278.
Ito, T.A., Miller, N., Pollock, V.E., 1996. Alcohol and aggression: a meta-analysis on the moderating effects of inhibitory cues, triggering events, and self-focused attention. Psychological Bulletin 120, 60–82.
Kaufman Kantor, G., Straus, M.A., 1990. The "Drunken Bum" Theory of Wife Beating. In: Straus, M.A., Gelles, R.J. (Eds.), Physical Violence in American Families: Risk Factors and Adaptions to Violence in 8,145 Families. Transaction, New Brunswick, NJ, pp. 203–224.
Leonard, K.E., 2008. The role of drinking patterns and acute intoxication in violent interpersonal behaviors. In: Alcohol and violence: Exploring Patterns of Violence. International Center for Alcohol Policies, Washington, DC, pp. 29–55.
Pernanen, K., 1991. Alcohol in Human Violence. Guilford, New York.
Quigley, B.M., Leonard, K.E., Collins, R.L., 2003. Characteristics of violent bars and bar patrons. Journal of Studies on Alcohol 64, 765–772.
Quigley, B.M., Leonard, K.E., 2005. Alcohol use and violence among young adults. Alcohol Research and Health 28, 191–194.
Quigley, B.M., Leonard, K.E., 2006. Alcohol expectancies and intoxicated aggression. Aggression and Violent Behavior 11, 484–496.
Steele, C.M., Josephs, R.A., 1990. Alcohol myopia: its prized and dangerous effects. American Psychologist 45, 921–933.
Sabra, B., Muller, D., Begue, L., Bushman, B.J., Delmas, F., 2010. Automatic effects of alcohol and aggressive cues on aggressive thoughts and behaviors. Personality and Social Psychology Bulletin 36, 1052–1057.
Taylor, S.P., Leonard, K.E., 1983. Alcohol and Human Physical Aggression. In: Green, R.G., Donnerstein, E.I. (Eds.), Aggression: Theoretical and Empirical Reviews. Academic Press, New York, pp. 77–101.

CHAPTER

64

Alcohol and Sexual Violence

Kelly C. Davis, Debra Kaysen, Amanda K. Gilmore, Trevor J. Schraufnagel
University of Washington, Seattle, WA, USA

OUTLINE

OVERVIEW

Sexual violence involves sexual contact that occurs without a person's consent. Anyone – men and women, boys and girls of all ages and backgrounds – can become a victim of a sexual violence, which may be perpetrated by partners, family members, friends, acquaintances, or strangers. This chapter focuses on the role of alcohol in situations involving sexual violence. We detail the definitions of various forms of sexual violence, provide prevalence rates of sexual violence and its co-occurrence with alcohol consumption, review the types of research conducted regarding alcohol and sexual violence, describe the alcohol-related pathways to both perpetration and victimization, and discuss efforts to prevent alcohol-involved sexual violence.

Definitions

A variety of terms have been used to describe nonconsensual sexual behavior, including sexual assault, sexual violence, sexual aggression, sexual abuse, sexual coercion, attempted rape, and completed rape. The terms *sexual violence* and *sexual assault* refer to any type of sexual contact that is not freely consented to by one of the persons involved. Thus, both of these terms have a fairly broad definition that includes a wide scope of behaviors ranging from nonconsensual sexual kissing or touching to nonconsensual oral, anal, or vaginal sex, and these general terms are used interchangeably throughout this chapter. The term *rape* refers specifically to a completed nonconsensual sex act in which the perpetrator penetrates the victim's vagina, anus, or

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00064-4

mouth with a penis, hand, finger, or other object. *Child sexual abuse* incorporates any nonconsensual sexual contact of a minor or where there is sexual contact between a minor and someone older than them, regardless of whether consent was obtained.

Consent involves actual words or actions indicating freely given agreement to engage in sexual activity. If one's "agreement" is in any way pressured, coerced, or forced, the individual did not truly consent to the sexual activity. Additionally, one's ability to give consent may be compromised by a number of factors including drug or alcohol intoxication (voluntary or forced), age (being a minor), or physical or mental disability. For example, a person who has passed out due to alcohol intoxication is not capable of consenting to sexual activity; thus, any sexual activity that occurs with an individual who is unconscious due to alcohol intoxication is considered sexual assault. Although clear in this example, many instances of sexual violence involve lesser amounts of alcohol intoxication which can lead to uncertainty about the provision of consent. Unlike drunk driving, there is no established blood alcohol content (BAC) for assessing one's capability of providing sexual consent. Thus, distinguishing intoxicated sex from rape creates challenges within forensic, clinical, and research domains. Importantly though, the presence of alcohol in sexually violent situations should in no way be used to justify, excuse, or minimize the occurrence of sexual assault or its effects.

PREVALENCE OF SEXUAL VIOLENCE

Sexual violence is present across the globe in all social, economic, ethnic, racial, religious, and age groups. According to the 2006 National Violence against Women Survey (NVAWS), over 300 000 women and almost 93 000 men are raped in the United States each year, and one of every six women and one in 33 men in the United States has been raped at some time in her or his life. Studies estimate that up to 80% of women experience sexual assault, including unwanted verbal or physical coercion. More than half of female sexual assault victims experienced their first sexual assault while under the age of 18.

While both women and men can be perpetrators and victims of sexual assault, by far the most typical sexual assault scenario involves a female victim and a male perpetrator. For example, of the sexual assault victims in the 2006 National Crime Victimization Survey (NCVS), 89.3% were female compared with 10.7% male. Because of this consistent pattern of victimization, the preponderance of research on sexual aggression has examined male-on-female sexual assault. As such, the current understanding of other patterns of assault (i.e. female-on-male or same-sex sexual assaults) remains extremely limited, and future research into such incidents (alcohol-involved or otherwise) is needed. However, for this chapter we primarily reference sexual assaults involving a male perpetrator and a female victim as this is the modal – and most studied – form of sexual violence.

Research indicates that most sexual assaults are committed by someone known to the victim. The NCVS found that approximately 62% of the sexual assaults reported in 2006 were perpetrated by someone known to the victim, while 38% were perpetrated by strangers. Rapes committed by strangers occurred in only 16.7% of the female victim cases reported in the NVAWS.

Prevalence of Alcohol-Involved Sexual Violence

One contributing factor to sexual assault risk is alcohol. Event-level survey research has consistently found that a majority (55–74%) of acquaintance sexual assault incidents involves alcohol use by male perpetrators, female victims, or both. In toxicology analyses of blood and urine samples from female sexual assault victims who reported their crime to the police, 43% of the samples had ethanol present, and 12% had ethanol and at least one other drug. Sexual assaults that involve alcohol consumption can differ in important ways from those that occur when both the victim and the offender are sober. For example, completed rapes are more likely when both victim and offender have been drinking than when both are sober. Moreover, there may be differences between assaults where the victim could not consent because of her level of intoxication versus those where she was given alcohol intentionally to facilitate sexual assault.

RESEARCH ON ALCOHOL AND SEXUAL VIOLENCE

Research regarding the relationship between alcohol and sexual violence has employed a variety of methods for examining the characteristics of alcohol-involved sexual assault, as well as the distal and proximal factors that may predict alcohol-involved perpetration and victimization. Survey studies investigating global correlations and situational covariation consistently report a positive association between alcohol use and sexual violence. While these nonexperimental findings point to the possibility of a relationship between alcohol use and sexual violence, these data are primarily cross-sectional and correlational, thereby precluding the ability to establish a causal stream of events.

Experimental investigations of alcohol's influence on sexual violence corroborate the relationship between

alcohol consumption and sexual violence. Moreover, controlled laboratory conditions allow for establishing a causal connection between these variables. Using either written, audio, or video-based hypothetical sexual encounters, laboratory studies have consistently found that after alcohol is administered to participants, intoxicated men report greater sexually aggressive intentions toward a hypothetical female partner than their sober counterparts, while intoxicated women are less perceptive of sexual assault risk cues. The experimental method affords several advantages over other methods, including the ability to test causal relationships and theoretical mechanisms of causation, tight control of important variables (e.g. alcohol dose, expectancy set, BAC, and limb of the blood alcohol curve), and isolation of in-the-moment moderating and mediating variables.

Focus groups and individual interviews have also been utilized to gather qualitative information regarding the relationship between alcohol and sexual violence. When interviewed, sexually aggressive men often cite alcohol consumption as a causal factor in the occurrence of rape. Approximately two-thirds of the perpetrators in one study viewed high levels of alcohol intoxication as a causal factor in the assaults. Studies of female focus groups indicate that women believe that alcohol intoxication decreases their ability to perceive sexual assault risk accurately and to engage in effective protective strategies. Much like surveys, these methods do not enable investigators to determine if the relationship between alcohol consumption and sexual violence is a causal one. However, these methods do allow researchers to capture a rich portrait of real-world sexual violence events.

CHARACTERISTICS OF ALCOHOL-INVOLVED ASSAULTS

Research suggests that episodes of sexual aggression involving alcohol often possess different characteristics than sexual violence perpetrated in the absence of alcohol. Episodes of alcohol-involved sexual aggression may differ in terms of the relationship between the victim and perpetrator as well as the characteristics of the aggression itself (e.g. setting, perpetrator use of force, etc.). The post-assault experiences of the victim may also be related to whether alcohol was involved or not.

Victim and Perpetrator Relationship

Compared to sexual assaults without alcohol involvement, alcohol-involved sexual assaults are less likely to be perpetrated by a person well known to the victim, although not all studies support this assertion. In one study, female victims that were impaired or incapacitated during the assault were more likely to report that the perpetrator was an acquaintance and less likely to report that the perpetrator was a romantic partner. Other studies indicate that victim–perpetrator relationship may vary according to whether only the perpetrator was drinking, only the victim was drinking, or both were drinking. Thus, while alcohol does seem to be related to the degree of familiarity between perpetrator and victim, the available data suggest that additional insight may be gleaned by examining precisely who was consuming alcohol and what volume was consumed.

Setting

Unsurprisingly, sexual violence that co-occurs with alcohol use is likely to occur in social settings or in venues that serve alcohol such as bars and nightclubs. One study found that alcohol was involved in only 6% of sexual assaults that occurred in a victim's home, whereas alcohol was involved 29% of the time if the assault occurred at a party or at another person's home. Women in these settings may be at increased risk of sexual violence due to their own drinking or the perpetrator's drinking, greater exposure to an increased number of male perpetrators, and/or men in these situations may be more likely to misconstrue women's sexual intent.

Level of Force

Whether or not alcohol was consumed prior to a sexual assault appears to be associated with the amount of force used by a perpetrator, but the nature of this relationship is affected by factors such as who is drinking and how much alcohol has been consumed. Current research generally suggests that perpetrator drinking is associated with increased use of force and greater victim injury. However, often both the perpetrator and victim have been drinking, and research suggests that victim alcohol use may temper perpetrator use of force. Indeed, several studies indicate that sexual assaults in which only the perpetrator was drinking were more violent, resulted in more injury, and inspired more fear in the victim than episodes in which both parties had consumed alcohol. In sum, while research suggests that alcohol consumption and the use of force during sexual assaults are not independent of one another, the relationship is not simple and may be influenced by how much alcohol was consumed and by whom was it consumed.

Assault Disclosure and Response to Victimization

Research suggests that victims of alcohol-related sexual violence blame themselves more for the assault

and receive less favorable social responses to assault disclosure than do victims of assault in which alcohol was not involved. Both of these factors may decrease the likelihood that the assault is disclosed or may affect to whom the assault is disclosed. However, published research findings are mixed. One study reported that when compared to non-impaired victims, victims that were impaired or intoxicated prior to their assault reported more self-blame but did not receive more frequent negative responses from others. However, impaired and intoxicated victims were more likely to disclose to informal sources of support (e.g. friends and family members) but not to formal sources of support such as the police or treatment providers. Another study found that self-blame was increased if the victim had consumed alcohol beforehand and that self-blame was the lowest if the perpetrator only had been drinking. Further, perpetrator-only drinking was also associated with more frequent disclosure and more positive and helpful social responses compared to assaults in which both parties were drinking and compared to non-alcohol-involved assaults (although greater assault severity may also have contributed to these findings).

ALCOHOL-RELATED PATHWAYS TO PERPETRATION AND VICTIMIZATION

Investigations of alcohol-involved sexual assault have relied upon two frames for conceptualizing how alcohol might foster sexual assault perpetration and victimization. The "distal" frame focuses on explaining factors associated with the likelihood that someone would be in a drinking situation where there is the potential of perpetrating, or being victimized by, sexual assault. Survey research has effectively identified important alcohol-related factors in this frame, including drinking habits and alcohol expectancies. The "proximal" frame focuses on explaining – once the person is in the situation – how alcohol can lead to a sexually violent outcome. Factors assessed in the proximal frame include acute alcohol intoxication and in-the-moment processes and reactions that may facilitate sexual assault. Experimental research has established that intoxication-related cognitive impairment, reduced behavioral inhibition, and sexual arousal may function as important factors predictive of sexual assault perpetration, while reduced risk perception and physical incapacitation may facilitate alcohol-related sexual victimization.

Distal Factors

Sexual Assault Perpetration and Alcohol Use

Research indicates that men who perpetrate sexual aggression report higher levels of alcohol use than the general population. For example, college men who reported sexual assault perpetration are more likely than men who did not report perpetration to meet the diagnostic criteria for alcohol abuse or dependence. Additionally, incarcerated sexual offenders are more likely to report alcohol problems than are incarcerated nonsexual offenders. Sexual assault severity has also been related to alcohol consumption such that the more an individual tends to drink, the more severe his sexual assault perpetration. It is critical to note, however, that analysis of these associations at the global level does not address possible third variables, such as sensation seeking or psychopathology, which may underlie the relationship between alcohol use and sexual assault perpetration.

Sexual Victimization and Alcohol Use

Greater alcohol consumption, more negative consequences associated with drinking, higher endorsement of problem drinking, and higher rates of diagnoses of alcohol use disorders have all been associated with sexual victimization. In general, higher alcohol use and problems have been demonstrated in treatment seeking, college student, and community samples. The association between sexual victimization and higher alcohol use and problems has been documented for heterosexual, lesbian, and bisexual women, as well as with gay men.

The majority of the research conducted to date on victimization and drinking outcomes has been conducted cross-sectionally comparing sexual victimized to non-victimized women or comparing women based on the level of severity of the sexual assault. Causal relationships between drinking and sexual assault have not yet been disentangled, in part because of the relatively low number of prospective studies. Several longitudinal studies have found a relationship between sexual assault and risk of problematic drinking later in time, although findings are mixed. Adolescents with histories of child sexual abuse have been found to be at higher risk for substance use outcomes including more problematic drinking. However, longitudinal studies conducted with adult victims of sexual assault have had more mixed findings, with some studies finding sexual assault predicts future high-risk drinking and other studies failing to find this relationship.

More consistently, alcohol use has been found to increase risk of later victimization. In general, women who drink more are at increased risk for sexual victimization. Increased substance use has been found to mediate the relationship between child sexual abuse and college revictimization. Binge drinking (defined as consuming either four standard drinks (for women) or five standard drinks (for men) in 2 h or less) and heavier drinking both appear to increase risk of later

victimization. Moreover, novice drinkers who experience a heavy drinking episode appear to be at especially high risk for sexual victimization.

Drinking to Cope with Prior Victimization

As noted above, women with a sexual assault history (including women with a history of child sexual abuse or adult sexual assault) drink more than women with no sexual assault history. This difference in drinking habits appears to put these women at higher risk for sexual revictimization. For example, alcohol consumption mediates the relationship between sexual assault and revictimization such that sexual assault history increases risk for heavy alcohol consumption, which reciprocally increases risk for additional sexual assault. Increased alcohol consumption after a sexual assault or after child sexual abuse may be due to drinking as a means of coping with distress from the assault. Women with a history of child sexual abuse or adult sexual assault are more likely to endorse that they use alcohol to cope than women with no sexual assault history. Coping motives, or endorsing the use of alcohol to regulate negative emotions, were also found to mediate the relationship between child maltreatment and negative consequences of drinking in both male and female college students and in a separate study were found to mediate the relationship between child sexual abuse and alcohol problems in adult women. It was also found in this study that sexual revictimization was predicted by the number of maladaptive coping strategies used, including the use of drugs or alcohol to cope. However, little research has been devoted to examining the relationship between drinking motives and sexual assault history; therefore, it is difficult to make generalizations.

Alcohol Expectancies

An alcohol expectancy is defined as one's belief about how people (or oneself) will behave while under the influence of alcohol. Alcohol expectancy theory contends that alcohol consumption may influence behavior through individuals' expectations about alcohol's emotional, physiological, and behavioral effects, which are often influenced by overarching cultural beliefs regarding alcohol-related outcomes. Alcohol expectancy theory posits that alcohol consumption can result in an increased likelihood of sexual assault outcomes due to commonly held beliefs in our society that alcohol increases sexual arousal, sexual behavior, sexual enhancement, and sexual aggression. Thus, sexual behaviors, including those that are coercive or aggressive, are viewed as more normative when performed under the influence of alcohol and may in turn occur more frequently while intoxicated.

Regarding sexual assault victimization, alcohol expectancy theory stipulates that positive alcohol expectancies may lead to increased drinking, which may be predictive of a higher risk of experiencing sexual assault. Research has documented global differences in alcohol expectancies based on sexual assault history such that women with a sexual assault history endorse more positive beliefs about what alcohol will do as compared to women with no sexual assault history. Studies examining the relation between alcohol expectancies and sexual assault have used various measures to investigate alcohol expectancies and frequently vary in how they define "sexual assault" (e.g. using different assault severity ratings or comparing women based on the presence or absence of substance use prior to the assault). However, despite these differences, there have been consistent findings that women with a sexual assault history report greater expectations that alcohol improves relaxation, enhances sexuality, and increases sexual disinhibition compared to women without a sexual assault history.

Experimental balanced placebo designs have enabled investigators to disentangle alcohol expectancy effects from physiological effects by independently manipulating alcohol content and alcohol expectancy set (the belief that one has or has not consumed alcohol) through the use of placebos, in which a study participant is told that she is consuming an alcoholic beverage when, in fact, the beverage is nonalcoholic. One study found that women who expected to receive, but did not receive, alcohol and believed that alcohol enhances sexual behavior generated less effective resistance responses to an audio-taped portrayal of a sexual assault relative to control participants, while another study documented that women who expected alcohol rated a "risky" man more positively than controls but not as much as intoxicated women did. However, a third study found no effect of expectancy set on women's sexual assault risk appraisals. It appears that beliefs about alcohol's effects may be an important element in women's resistance; however, their role in predicting risk appraisals is less clear.

In terms of sexual assault perpetration, sexually aggressive men, especially those involved in alcohol-related assaults, report a greater belief in alcohol's enhancement of men and women's sexuality as compared to men without a history of sexual assault. Further, stronger alcohol expectancies regarding sex drive distinguish alcohol-involved perpetrators from both non-alcohol-involved perpetrators and non-perpetrators. Thus, beliefs that alcohol enhances sexual responding can increase men's sexual arousal and their perceptions of women's sexual arousal and may, through a form of self-fulfilling prophecy, result in increased sexual aggression likelihood.

This idea is supported by several alcohol administration studies that have assessed individually held alcohol

expectancies regarding sexual aggression before the administration of alcohol to determine whether the strength with which individuals hold such beliefs influences their responses. In one such study, men with stronger sex-related alcohol expectancies reported a greater likelihood of sexual aggression as mediated by greater sexual and affective responding to an eroticized rape depiction. Similarly, men with stronger beliefs about women's sexual vulnerability when drinking reported greater sexual arousal to an eroticized rape depiction, which in turn predicted greater estimates of their sexual aggression likelihood.

Although many experimental balanced placebo studies have found significant effects for expectancy set (i.e. placebo conditions) on dependent sexual aggression measures such as the amount of time to make a decision to stop unwanted sexual advances (i.e. decision latency), these effects are smaller than those for actual alcohol consumption. One study found that alcohol expectancy set had indirect effects on men's estimations of their sexual aggression likelihood. However, a similar number of studies have not found expectancy set effects for sexual aggression-related dependent measures. Thus, results regarding the influence of expectancy set manipulations on sexual aggression dependent measures are both inconsistent and relatively modest.

Alcohol expectancy theory contrasts with "deviance-disavowal," a theoretical conceptualization of alcohol's effects which suggests that alcohol consumption may contribute to deviant or anti-normative behavior by granting intoxicated individuals a respite from social norms, thus providing an excuse for formerly inexcusable behavior. The appeal of deviance-disavowal explanations for sexual aggression seems to be largely intuitive and not empirical, with one recent study finding that men who were intoxicated at the time of perpetration actually rated their offense as more serious and accepted more blame than did sober perpetrators – in direct contrast to what would be predicted within a deviance-disavowal framework.

Proximal Factors

Cognitive Impairment

According to alcohol myopia models, alcohol intoxication pharmacologically reduces cognitive processing ability, resulting in a narrowed attentional focus. The underlying notion is that as intoxication level increases a person becomes less able to process information in the environment. This cognitive impairment results in restricted focus on the most salient situational cues. Because sexual assault usually occurs after a man and a woman have been socializing, cues associated with having a good time or possibly developing a relationship may be more prominent than sexual assault cues. For men, this may mean that alcohol intoxication decreases their ability to process less salient verbal or physical resistance cues. For women, this may mean that alcohol intoxication decreases their ability to process sexual assault risk cues.

In reference to sexual assault risk appraisals, both experimental and survey research have found that alcohol intoxication decreases sexual assault risk perception; however, not all studies find support for this relationship. Most experimental and retrospective surveys of sexual assault victims have found that alcohol consumption decreases the ability to perceive risk cues, and one study showed that risk perception mediated the effect of alcohol on resistance. However, a qualitative analysis showed that intoxicated and sober women did not differ in their risk recognition. Therefore, the relationship between alcohol intoxication and sexual assault risk perception is not yet clearly understood.

Regarding perceptions of the victim's actual sexual interest and willingness, research indicates that alcohol increases men's likelihood of misperceiving their partner's true desires. Experimental investigations have shown that intoxicated participants tend to perceive ambiguous cues as indicative of sexual interest, more so than their nondrinking counterparts. A survey of 814 undergraduate men supported these findings: Men with higher overall alcohol consumption also reported a higher frequency of misinterpreting women's sexual intentions. Moreover, intoxicated men perceived female acquaintance rape victims as more sexually aroused and enjoying themselves more during the assault than sober men. One study found that when presented with a dating situation in which unwillingness cues were strong and clear, alcohol consumption did not influence men's ratings of acceptability of forced sex. However, when sexual interest cues were salient but unwillingness cues were not, men who consumed a moderate dose of alcohol (BAC 0.065%) rated forced sex as more acceptable than did controls. These findings support an alcohol myopia explanation in that intoxicated men had a greater focus on female sexual arousal rather than on less salient unwillingness cues.

Perpetrator Behavioral Disinhibition

Another pharmacologically based theory of alcohol's effects on aggression suggests that alcohol may directly (versus myopia theory's indirect-effect prediction) allow for aggression by increasing behavioral disinhibition. That is, alcohol intoxication may impair one's typical restraint of socially unacceptable or otherwise problematic impulses or behaviors – a supposition with recent support in the cognitive neuroscience literature. Two studies have compared a myopia model to a behavioral disinhibition model in the realm of alcohol and sexual assault; however, the first supported a myopia

perspective and the second supported disinhibition. Given the equivocal findings and that both of these studies have been criticized for their methodological limitations, further research is needed to evaluate the nature of alcohol's pharmacological effects on sexually aggressive behavior.

Perpetrator Sexual Arousal

Because both male perpetrators and female victims indicate that sexual assaults which occur in dating situations are typically preceded by consensual sexual activities, perpetrator sexual arousal may play a key role in predicting sexual assault likelihood, particularly among acquaintances. In support of this idea, one experiment manipulated men's level of sexual arousal and found that men in the high arousal condition reported a greater likelihood of engaging in acquaintance sexual aggression in a hypothetical situation than did men in the non-arousal condition. Notably, alcohol intoxication, or the expectancy thereof, can facilitate men's sexual arousal and thereby increase the likelihood of sexual aggression perpetration. In support of this, a laboratory study reported that when reading an eroticized rape depiction, intoxicated men reported greater sexual arousal than did sober men, which in turn increased their sexual aggression intentions.

Victim Resistance, Passivity, and Incapacitation

Survey studies have documented that women who were intoxicated at the time of a sexual assault generally report that they resisted the assault less assertively than did women who were sober at the time of the assault. Similarly, acute alcohol intoxication studies have found that, compared to sober women, intoxicated women report less resistance of hypothetical unwanted sexual advances. For example, several studies have found that intoxicated women were more likely to respond passively or politely compared to sober women. Other studies have found that women's sexual assault risk perception or other appraisals of the situation partially mediated the relationship between alcohol and resistance intentions. Thus, because assertive resistance can often thwart a sexual assault, intoxicated women may be more likely to experience a completed rape than sober women in that they are more likely to engage in generally ineffective forms of resistance. Moreover, heavily intoxicated women who are passed out or otherwise physically or mentally incapacitated are obviously unable to engage in sexual assault resistance of any type.

Synergy of Distal and Proximal Factors

Research suggests that sexual assault history and alcohol intoxication may synergistically increase the likelihood of sexual assault. In women, a history of childhood or adult sexual assault has been associated with adult revictimization. This may occur because previously victimized women do not perceive sexual threat cues or may not effectively resist new assaults. Childhood sexual abuse has been related to increased immobility and passivity in adolescent/adult sexual assaults, while prior adolescent/adult sexual assault has also been associated with heightened non-forceful resistance and lowered assertive resistance.

To the extent that alcohol intoxication impairs women's ability to assess risk accurately and resist effectively, it may interact synergistically with women's pre-existing tendencies regarding risk perception and resistance (which may themselves be associated with prior victimization). For example, in a study using latent profile analysis of women's actual experiences with sexual assault, women with the highest level of alcohol consumption during the assault and high levels of prior victimization felt more powerless, recognized that alcohol made them vulnerable and impeded their resistance, and engaged in low levels of resistance. Women who had consumed alcohol but were low in prior victimization reported that alcohol made them vulnerable, that alcohol was a barrier to their self-defense, and resisted through diplomatic tactics. These profiles show how alcohol intoxication and victimization history can both individually and concomitantly relate to specific types of risk appraisals and resistance responses.

For men, distorted cognitions, particularly in relation to the appraisal of women's sexual interest and willingness, have been theorized to act as key precursors to sexual aggression in that they allow sexually aggressive men to ignore a woman's refusal cues that would typically inhibit the continuation of sexual pursuit. Research has corroborated this notion that sexually aggressive men's cognitive appraisals differ from those of nonsexually aggressive men.

The cognitive impairment that results from alcohol intoxication may further exacerbate these cognitive distortions. That is, men with a proclivity toward sexual aggression may be most likely to commit assault when intoxicated. A laboratory experiment found that intoxicated sexually coercive men took significantly longer than sober sexually coercive men to determine when a man should stop unwanted sexual advances. Moreover, sexually aggressive men often cite alcohol's effects on their perceptions as a causal factor in the occurrence of rape. In interviews with self-disclosed date rapists, almost one-half commented that alcohol intoxication increased perceptions of their partner's sexual willingness. Alcohol expectancies may further moderate these proximal effects of intoxication.

SEQUELAE OF ALCOHOL-RELATED VICTIMIZATION

There is a large body of literature detailing negative long-term effects associated with sexual victimization, including increased rates of major depression (isolated or recurrent episodes of severely depressed mood lasting for at least two weeks), posttraumatic stress disorder (PTSD), suicidal behavior, substance use disorders, and risk of revictimization. However, only recently have researchers begun to examine whether there are differences in psychological effects between alcohol-involved versus non-alcohol-involved assaults.

Elevated risk of PTSD has been associated with both child sexual abuse and adult sexual victimization. PTSD involves a constellation of symptoms that develop after exposure to a life-threatening traumatic event and may include persistent re-experiencing of the traumatic event, avoidance of trauma-related stimuli, emotional numbing, and increased physiological arousal and reactivity. Completed rape is associated with one of the highest rates of PTSD across traumatic events. Studies conducted to date examining the relationship between alcohol-involved assault and PTSD yield contradictory results. Some studies have found lower rates of PTSD associated with alcohol use, whereas other studies have found no differences in PTSD rates between alcohol-involved assaults and other assaults. However, in these studies, PTSD has often been assessed retrospectively, making it impossible to examine alcohol's role in the development of PTSD symptoms. Based on the few prospective and longitudinal studies, alcohol may alter the trajectory of development of PTSD over time. Alcohol-involved assaults have been associated with lower initial PTSD symptom severity. This could be explained by alcohol's pharmacological effects on stress and on memory. Acute alcohol administration can dampen stress responses. Alcohol also has been shown to reduce acquisition of fear memories. For example, in a PTSD analog study, higher doses of alcohol were associated with fewer intrusive memories of a stressful film a week later than what was reported by participants who had not consumed alcohol or who had consumed low doses. These differences also may be due to the assault characteristics mentioned above that may differ between non-alcohol-involved assaults including degree of force during the assault, relationship with the perpetrator, and the extent to which the victim thought they would be injured or killed during the assault. Similarly, some studies have found no differences in rates of depression between alcohol-involved and non-alcohol-involved assaults, whereas other studies have found lower rates of depression in the alcohol-involved assaults. As this body of literature is still very small, it is too early to be able to determine the role of alcohol at the time of the assault on the long-term risk of depression. It is important to note that alcohol-involved assaults are still associated with PTSD and depression, even in those studies that have found higher rates of those disorders associated with forcible rapes.

There do appear to be more consistent differences between alcohol-involved and non-alcohol-involved assaults in substance use outcomes. Women who have experienced alcohol-involved assaults are at higher risk for later problems with alcohol use including higher consumption, higher rates of binge drinking, and higher rates of substance abuse. However, the majority of these studies have only looked at cross-sectional relationships. As alcohol use itself increases the risk of experiencing alcohol-related sexual assault, it may be that these findings predominantly reflect risk factors for alcohol-related assaults rather than reflecting consequences of experiencing an assault per se.

PREVENTION OF ALCOHOL-RELATED ASSAULTS

To date, the vast majority of rape prevention programs have been directed at women, as women are predominantly the victims of sexual victimization. Despite a large number of rape prevention programs, there is little research suggesting that these programs are effective in reducing sexual assaults. In general, these programs focus on changing attitudes, such as decreasing myths about rape, increasing women's self-efficacy, or increasing women's ability to use self-defense. They often include education about rape definitions and prevalence, relate information about sexually coercive behavior, discuss high-risk situations and how to increase safety while dating, and provide resources for victims. There is a lack of well-controlled randomized clinical trials to evaluate these programs, despite how widely they are implemented. The studies that have been conducted evaluating rape prevention programs for victims have generally failed to demonstrate changes in the rates of victimization or in those behaviors that are thought to increase risk of being assaulted. Interestingly, one of the few prevention programs to show an effect in decreasing victimization did not find that the prevention program worked for women who used alcohol, despite their overall elevated risk for victimization. Despite the link between alcohol use and increased risk of victimization, few prevention programs address alcohol use as a means of reducing risk of victimization, nor have studies examining prevention of high-risk drinking in college students

addressed whether these general prevention programs may work to reduce risk of victimization. Early findings from one study suggest that reducing binge drinking prior to entry into college may be a means by which to reduce victimization, especially during the most high-risk period for sexual assault. Although interventions regarding alcohol use are unlikely to address all types of sexual victimization, it is promising as a means to address alcohol-related sexual assaults.

While the preponderance of sexual assault prevention work is conducted with females and aims to reduce vulnerability to violence, research has also examined the effect of programs designed to enlist male participation in assault prevention. These studies largely show that the attitudes, knowledge, and empathy (which have been linked to violence perpetration) of men can be modified via various programs. Surprisingly, only one study has reported findings regarding alcohol-related aggressive behaviors and attitudes following an intervention. The program examined in this study (The Men's Program) included an increased emphasis on alcohol based on qualitative data from an earlier project. By bolstering men's self-concept as helpers having the potential to intervene, The Men's Program resulted in men reporting greater concern over sex while intoxicated as well as the role alcohol may play in comprising a woman's ability to consent to sexual activity. Moreover, these qualitative data also evidenced behavioral changes. For instance, participants noted intervening to prevent their friends from engaging in alcohol-related sexual activity. Men also reported abstaining from sexual activity if either they or a prospective partner were intoxicated.

SUMMARY

Sexual violence is a major public health risk. Because the majority of sexual assault incidents involve alcohol consumption by either the perpetrator, the victim, or both the perpetrator and victim, scientists have labored to understand the ways in which alcohol-related factors may influence the occurrence, phenomenology, severity, and sequelae of sexual assault. These empirical investigations have yielded important information about the alcohol-related pathways toward sexual violence. First, distal factors such as typical drinking habits and alcohol expectancies may increase the likelihood of a particular individual being in a drinking situation that involves an elevated risk of sexual violence. Second, proximal factors like alcohol-induced cognitive and physical impairment may influence one's actions before or during the event in ways that increase the potential for sexual violence. Phenomenological differences such as the prior relationship between victim and perpetrator and the degree of force used during an assaultive incident also tend to differ based on whether alcohol was involved. Additionally, the sequelae of alcohol-involved sexual assaults, such as PTSD symptom development and substance use outcomes, may differ in important ways from assaults that do not involve alcohol. As research continues to identify and parse the psychological and physiological mechanisms through which alcohol influences both the occurrence and consequences of sexual violence, programs addressing the prevention of perpetration and reduction of victimization will necessarily require continued modification and rigorous evaluation. Finally, although the results of extant research clearly indicate that alcohol is a risk factor for both sexual assault victimization and perpetration, alcohol consumption should neither diminish perpetrator accountability nor increase victim culpability. As with other crimes, while alcohol may facilitate sexual violence, it does not excuse it.

SEE ALSO

Binge Drinking, Impulsivity, Disinhibition, and Risk Taking in Addiction, Substance Induced Myopia, Developmental Risk Taking and the Natural History of Alcohol and Drug Use among Youth, Drinking Patterns, Alcohol Consumption, and Aggressive Behavior, Alcohol, Sexual Risk Taking, and Sexually Transmitted Infections

List of Abbreviations

BAC	blood alcohol content
NCVS	National Crime Victimization Survey
NVAWS	National Violence against Women Survey
PTSD	posttraumatic stress disorder

Further Reading

Abbey, A., Zawacki, T., Buck, P.O., Clinton, A.M., McAuslan, P., 2004. Sexual assault and alcohol consumption: what do we know about their relationship and what types of research are still needed? Aggression and Violent Behavior 9, 271–303.

Brecklin, L.R., Ullman, S.E., 2010. The roles of victim and offender substance use in sexual assault outcomes. Journal of Interpersonal Violence 25, 1503–1522.

Foubert, J.D., Godin, E.E., Tatum, J.L., 2010. In their own words: Sophomore college men describe attitude and behavior changes resulting from a rape prevention program 2 years after their participation. Journal of Interpersonal Violence 25, 2237–2257.

Giancola, P.R., Josephs, R.A., Parrott, D.J., Duke, A.A., 2010. Alcohol myopia revisited: clarifying aggression and other acts of disinhibition through a distorted lens. Perspectives on Psychological Science 5, 265–278.

Testa, M., 2002. The impact of men's alcohol consumption on perpetration of sexual aggression. Clinical Psychology Review 22, 1239–1263.

Testa, M., Livingston, J.A., 2009. Alcohol consumption and women's vulnerability to sexual victimization: can reducing women's drinking prevent rape? Substance Use & Misuse 44, 1349–1376.

Ullman, S.E., 2003. A critical review of field studies on the link of alcohol and adult sexual assault in women. Aggression and Violent Behavior 8, 471–486.

Ullman, S.E., Najdowski, C.J., 2010. Understanding alcohol-related sexual assaults: characteristics and consequences. Violence and Victims 25, 29–44.

Relevant Websites

http://www.niaaa.nih.gov/ – National Institute on Alcohol Abuse and Alcoholism.

http://www.vawnet.org/ – National Online Resource Center on Violence against Women.

http://www.musc.edu/vawprevention/ – National Violence against Women Prevention Research Center.

http://www.ovw.usdoj.gov/ – Office on Violence against Women at the U.S. Department of Justice.

http://www.rainn.org/index.php – Rape, Abuse, & Incest National Network.

CHAPTER

65

Alcohol, Sexual Risk Taking, and Sexually Transmitted Infections

*Melissa A. Lewis**, *Megan E. Patrick*[$]*, *Keith J. Horvath***, *Dana Litt**

*University of Washington, Seattle, WA, USA [$]University of Michigan, Ann Arbor, MI, USA

**University of Minnesota, Minneapolis, MN, USA

OUTLINE

Research has demonstrated that a vast majority of individuals in the United States report using alcohol in the last month. Moreover, roughly 75% of the alcohol consumed in the United States is in the form of heavy-episodic drinking (often referred to as binge drinking), which is typically defined as having at least four or five drinks on a single occasion during a specified time period for females and males, respectively. The proportion of current drinkers that engage in heavy-episodic drinking is highest among 18- to 20-year-olds, with roughly 90% of all alcohol consumed in the form of heavy-episodic drinking. In addition to young adults, research indicates that adolescents (ages 12–18) engage in alcohol use. In 2005, about 10.8 million persons ages 12–18 (28.2% of this age group) reported drinking alcohol in the past month and nearly 7.2 million (18.8%) were classified as heavy-episodic drinkers. Individuals, and adolescents and young adults in particular, experience a range of alcohol-related negative outcomes such as poor class and/or work attendance, damaging property, hangovers, trouble with authorities, injuries, unprotected sex, sexual assault, and death. Thus, research demonstrates that high-risk drinking is problematic in the United States, particularly among adolescents and young adults.

AT-RISK POPULATIONS

Adolescents and Young Adults

The Centers for Disease Control and Prevention (CDC) estimates that more than one million people are living with human immunodeficiency virus (HIV) in the United States. Of particular concern, it is also estimated that one in five (21%) of those people living with HIV is unaware of their infection. In addition to HIV, sexually transmitted infections (STIs) remain a major public health challenge

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00065-6

in the United States. It is estimated that there are approximately 19 million new STIs each year. Almost half of these new STIs are among young people 15–24 years of age. The number of HIV and acquired immune deficiency syndrome (AIDS) cases that occurred among this same age group accounted for approximately 14% of all HIV/AIDS cases in 2006. Thus, similar to high-risk alcohol use, research demonstrates that STIs and HIV are a public health concern in the United States, particularly among adolescents and young adults. These findings also highlight the importance of examining any hindrances to condom use, such as alcohol consumption, as condom use is a reliable method of STI and HIV prevention for individuals who choose to engage in penetrative sex (i.e. vaginal or anal sex).

Sexual Minority Populations

In addition to adolescents and young adults, men who have sex with men (MSM) are also a group at risk for heavy alcohol use and STIs/HIV. According to the latest surveillance data from the CDC, MSM account for an estimated 61% of all new HIV infections in the United States in 2009, despite representing approximately 2% of the adult male population. A recent study by the CDC demonstrated that 19% of gay and bisexual men in 21 cities were infected with HIV, with nearly half of those infected unaware of their serostatus. For this reason, attention to the association between substance use – including alcohol – and sexual risk-taking among sexual minorities has primarily been on HIV risk behaviors among MSM, with relatively few studies examining the role of alcohol in STIs other than HIV or on sexual minority women or transgender populations.

It is important to recognize that our understanding of the associations between alcohol use/misuse and STI/HIV acquisition among sexual minority populations is hampered by a number of factors. Fully explicating alcohol use patterns among sexual minority persons and their associations with STI acquisition is complicated by limitations in sampling and assessment. Perhaps most problematic may be that it is not possible to determine the sampling frame for sexual minorities, and therefore obtaining a truly nationally representative sample of lesbian, gay, bisexual, transgender, queer, or questioning (LGBTQ) is not feasible (although we recognize that the term "transgender" reflects gender orientation, for brevity we refer to the spectrum of LGBTQ individuals as "sexual minorities" throughout this chapter). Even when questions about sexual orientation or same-sex behavior are included in national surveys, respondents may underreport on these items for fear of being "outed." Problems in sampling these hidden populations are reflected in the overrepresentation of white and highly educated respondents in studies of LGBTQ individuals. With respect to assessment, it is widely recognized that sexual orientation is multidimensional and includes identity (e.g. gay, lesbian, bisexual, or heterosexual), behavior (i.e. same- or other-sex sexual contact regardless of identity), and attraction (the degree of sexual affection toward same and/or opposite sex persons). Researchers have not established standard terms and definitions for assessing these dimensions of sexual orientation, and oftentimes these and other terms to capture sexual orientation are used interchangeably. Difficulties in establishing common definitions for sexual orientation mirror the lack of standardized measures in the assessment of alcohol consumption and time frame.

Alcohol use and alcohol-related problems among LGBTQ persons cannot be attributed to sexual minority status by itself, but rather emerge from a complex array of historical and social factors that stigmatize and marginalize these communities. Foremost, it must be emphasized that, in spite of stigmatizing experiences faced by LGBTQ individuals, most persons who identify as a sexual minority or engage in same-sex sexual behavior do not drink heavily or experience alcohol-related problems. Estimates of heavy drinking among gay men range from 14 to 19%, compared to 8 to 11% among heterosexual men. Similarly, studies comparing alcohol use patterns between lesbian and heterosexual women show more frequent alcohol use and higher rates of drinking to intoxication among lesbian women. However, other studies using probability samples of MSM show similar rates of heavy or frequent drinking to the general population.

ASSESSMENT OF ALCOHOL'S ROLE IN SEXUAL RISK TAKING AND STIs

Alcohol use before and during sex is common (i.e. one-fourth to a half of people reported doing so during the most recent sexual encounter). Research by Patrick and Maggs found that in a 14-day period, 40% of college students who engaged in sexual behavior reported drinking before or during at least one sexual encounter. Although alcohol has a pharmacological effect of impairing sexual performance, a strong cultural and popular belief that drinking increases the likelihood and/or pleasure of sexual experiences persists. At the same time, research has shown that individuals tend to overestimate how much alcohol they can consume without experiencing negative consequences, including unwanted sex. In other words, it seems that there is a prevailing cultural belief that alcohol facilitates sex without increasing risk for negative sexual consequences. In fact, college students reported experiencing

more positive consequences (e.g. enhancing sexual arousal) of sex if they had sex after drinking. However, this may result from self-fulfilling prophesies regarding the role of alcohol in sexual behavior rather than a true effect of alcohol consumption itself. In addition, the *negative* consequences (e.g. STI contraction, pregnancy) are not immediately evident and therefore cannot be reported accurately in short-term self-report surveys.

The particular intersection of alcohol use with sexual behaviors is not yet clearly understood. Although we know that individuals who drink more alcohol also tend to engage in more sexual behavior, the causal direction is not clear. In terms of describing potential causal mechanisms, there are three possibilities. First, there may be a causal effect from alcohol to sex because of the disinhibitory effects of drinking. For example, alcohol myopia theory suggests that the disinhibitory effects of alcohol allow immediate goals and impulses to influence behavior more strongly than long-term goals. In that case, sexual arousal may be the most salient impulse and lead to sexual behavior, and perhaps riskier sexual behavior, when individuals are intoxicated. Second, there may be a causal effect from sex to alcohol. That is, people may drink alcohol based on the expectation that it will make sex more likely. When individuals hold expectancies that alcohol will facilitate sexual behavior or feelings (e.g. sexual drive, sexual affect), they may be more likely to drink for these effects. Third, the association between alcohol and sex may be a spurious relationship, or the result of a common third variable, such as sensation seeking or spending more time in a setting that provides opportunity to engage in both behaviors (such as bars). The notion that alcohol consumption results in sexual risk taking (e.g. sex without a condom) is still being researched. In addition to these explanations, alcohol consumption might increase the biologic susceptibility to an STI, if exposed, which may occur through adverse effects on the immune system or other direct biologic changes.

The strength of the association between alcohol with sexual risk taking and risk for STIs appear to vary by whether alcohol and sexual behavior are measured globally (i.e. alcohol consumption in an extended time frame), at the situation level (e.g. alcohol use in the context of sexual activity), or at the event level (i.e. alcohol consumption at a specific sexual encounter). For these reasons, clearly identifying the relationship between alcohol use and sexual risk taking is a challenge.

Global-Level Research

A commonly used method to study the relationship between alcohol and sexual risk taking is global-level research. In most global-level research studies, participants report information about the quantity and frequency of their alcohol use over some extended period of time (e.g. past 3 months). Additionally, participants provide information about the number of times they engaged in specific sexual risk and protective behaviors during the same extended time period. These two pieces of information for alcohol and sexual behavior are then correlated to determine the strength of relationship between alcohol use and sexual risk-taking variables. Overall, a vast majority of global-level studies show that participants who use alcohol, and particularly those who use alcohol heavily, are more likely to report sexual risk-taking behaviors. For example, a national survey of college students found that heavy-episodic drinkers were nearly three times as likely to report having multiple sexual partners in the last month than were non-heavy-episodic drinkers. The same relationship also appears to hold for adolescents, ages 12–17, such that reporting engaging in heavy-episodic drinking has been linked with risky sexual behavior.

Another line of global association research looking more specifically at the relationship between alcohol use and STIs has found that problem drinking is associated with treatment for STIs and history of past infections. In particular, previous research indicates that young adults who report substance abuse are at greater risk for STIs. Based on a 2005 national survey of young adults ages 18–25, 3.9% who used both alcohol and illicit drugs and 3.1% who were heavy drinkers had an STI in the past year compared to 1.3% of those who did not use alcohol or illicit drugs. A meta-analysis found alcohol use increased the likelihood of incident HIV infection such that heavy-episodic drinkers had more than double the risk for acquiring HIV.

Arguably, when examining the relationship between alcohol use and STI/HIV, the most consistent finding is that heavy alcohol use is associated with HIV infection and unprotected anal intercourse (UAI), the sexual behavior most likely to transmit HIV and other STIs. Although the association between drug use and UAI has been clearly demonstrated, the degree to which alcohol use in low or moderate amounts is associated with increased risk taking or infection is less clear.

As noted earlier, studies of the association between alcohol consumption and sexual risk behavior or STI acquisition that focus on non-MSM sexual minority groups are relatively sparse. In one of the few studies of alcohol use and STI rates among lesbian women, Lindley and her colleagues found that women who self-identified as bisexual (9.1%) were more likely to report an STI in the past year than women who were unsure of their sexual orientation (6.3%), heterosexual (5.8%) or lesbian (2.3%). Slightly more women who identified as heterosexual (55.8%) or bisexual (55.5%) reported engaging in heavy-episodic drinking the last

time they "partied" than women who were unsure of their sexual orientation (49.8%) or were lesbian (48.3%). Although high rates of HIV (12–28% depending on method of assessing HIV status) have been found among male-to-female transgender groups, little is known about the contribution of alcohol use to sexual risk behavior and STI acquisition among this population. There is a demonstrable need for greater attention on the contribution of alcohol use to sexual risk behavior and STI acquisition among non-MSM sexual minorities.

These findings highlight the potential role that alcohol use plays in regard to STI/HIV risk for several high-risk populations. The results of these different lines of global association studies support the notion that alcohol and risky sexual behavior are linked. However, these types of global association studies do not provide information about the mechanisms that might link alcohol consumption and sexual risk beavior.

Event-Level Research

To address the limitations of global-level research, event-level methodologies have been developed to more adequately study the causal link between alcohol use and sexual behavior. Event-level studies can capture the dynamics and context of one or more particular sexual episodes rather than assessing general alcohol and sexual risk behavior patterns across a given time period. The primary advantage of event-level research is that detailed information is gathered regarding specific sexual events during which alcohol was consumed. This detailed information often includes, but is not limited to, whether a condom was used, amount of alcohol consumed at each event, as well as other factors (e.g. relationship status) that may exert influence on the relationship between alcohol use and sexual risk-behavior.

Event-level studies generally show that alcohol intoxication is associated with greater sexual risktaking depending on the outcome being evaluated. Overall, event-level studies yield inconsistent results when examining alcohol and condom use, with some studies finding that alcohol use leads to increased likelihood of using condoms, whereas others have found null or inverse results. A meta-analysis of event-level studies concluded that alcohol use is not necessarily associated with condom nonuse. Rather, other factors, such as the context, relationship between the participant and the sexual partner, and experience of the sexual partners are more influential. Morrison and colleagues found that of participants who reported sexual activity with and without prior drinking, 37% used condoms similarly whether drinking or not, 29% used condoms less frequently when drinking, and 34% used condoms more frequently when drinking. These idiosyncratic differences underscore the likely importance of mediating and moderating factors, such as alcohol expectancies, sensation seeking, relationship status, and gender, on the alcohol–condom use relationship.

Research has shown that the association between alcohol use and sexual protection behaviors is moderated by the nature of the relationship between sexual partners. Whether the partners have just met, are acquaintances, are occasional partners, or are in a committed relationship influences the decisions to have sex and to use a condom. A series of studies have examined the differential impact of relationship status, with a majority finding that more drinking before sex is more likely with casual partners, although more condom use is also more likely with casual partners.

In addition to partner relationship status, gender may be an important moderator of the event-level relationship between alcohol consumption and condom use, with men more likely to report condom use than women. This may be in part because different standards and social norms apply to men and women regarding condom use. Research indicates that men report less embarrassment when purchasing condoms, and women may receive (or at least perceive) negative evaluations for providing or carrying condoms. Moreover, because men ultimately have the decision of whether or not to wear a condom, women are at a disadvantage for successfully negotiating condom use. In event-level studies, findings are inconsistent regarding the role of gender as a moderator of the alcohol and condom use relationship. Some studies indicate a stronger relationship among women, others among men, and some report no gender differences.

Event-level studies on MSM populations show that MSM who have more than four alcoholic drinks prior to or during sexual activity are more likely to engage in risky sexual activity than those who consume fewer drinks. In contrast, event-level studies in which only the presence of alcohol before sexual activity was assessed have typically found little or no association with UAI. In sum, heavy alcohol use, but not consumption in lower amounts, may contribute to sexual risk taking and possible STI transmission among MSM.

INFLUENCES ON ALCOHOL AND SEXUAL RISK TAKING

Individual-Level Influences

A key to preventing alcohol use and related negative consequences in any population is gaining a better understanding of why individuals initiate and continue to engage in those behaviors. Individual-level influences include a range of genetic and biological factors, sociodemgraphic (e.g. age, gender, race, and ethnicity), and

alcohol use expectancies factors. Alcohol outcome expectancies vary across persons as a function of their learning histories and drinking experiences. Alcohol outcome expectancies also vary within person across time; as one has different experiences with or makes changes to their drinking over time, her or his alcohol outcome expectances may change.

Social Influences

Several commonly acknowledged social influences on the association between alcohol use and sexual behavior include the role of peers, family environment, modeling, and perceived social norms. These influences are thought to begin early in life and change across the life course. During adolescence, the role of peers becomes increasingly important as teenagers spend more time together, develop more intimate relationships, and begin experimenting in romantic relationships. In addition, the time between adolescence and young adulthood (sometimes referred to as emerging adulthood taking place between 18 and 25 years of age) is characterized as a period of exploration. Substance use peaks in the early twenties, and the majority of young adults (62% of males and 70% of females) are sexually active before age 18.

Influences Specific to Sexual Minority Populations

Although results vary across studies, research on patterns of alcohol consumption and problematic alcohol use suggests that rates of both alcohol use frequency and alcohol-related problems among LGBTQ communities may be higher than their heterosexual counterparts as a function of multiple interfacing influences. Although studies of alcohol use patterns in the general population overall show that alcohol consumption decreases with age, studies of lesbian and gay men show smaller declines in alcohol use and alcohol-related problems as they age. Gay men and lesbian women who hold expectancies that alcohol will reduce tension are more likely to drink alcohol, become intoxicated more frequently, and report alcohol-related problems than those who do not. However, by themselves, individual-level factors account for only a small proportion of the variance in alcohol use patterns and must be considered along with other factors.

Specific to the LGBTQ experience, "coming out" (i.e. disclosing one's LGBTQ identity) to family and peers as a member of a stigmatized group may be psychologically distressing, particularly when members of one's social network hold negative views toward sexual minorities. Although the association between psychological distress and increased alcohol use and problems remains tenuous, a study by McKirnan and Peterson found that nearly one quarter (22%) of gay men reported using alcohol at least half of the time to cope with personal stress. Because bars and drinking establishments traditionally have been one of the few places where LGBTQ individuals could socialize, peer norms around alcohol acceptance and use were widely adopted within these communities. Partner factors affecting alcohol use and consumption include relationship satisfaction and distress. The absence of institutional and family support for nonheterosexual partnerships may be associated with relationship disruption within LGBTQ partnerships, which has been shown to be associated with drinking-related problems. Although the association between intimate partner violence and alcohol consumption in LGBTQ communities is still unclear, some studies suggest that alcohol and other substance use often precedes violent events.

Social roles and expectations for lesbian women and gay men are significantly different than for heterosexual women and men. Most notably, fewer LGBTQ-identified persons have children than their heterosexual counterparts and until recently being legally married was not an option for most sexual minorities; both having children and being in a legally recognized marriage appear to be protective factors against alcohol-related problems. Particularly for lesbian women, having employment outside the home and employment in male dominated occupations are associated with more alcohol consumption. Both gay men and lesbian women who experience workplace discrimination and harassment are more likely to experience psychological distress and use alcohol as a coping strategy. Well-documented environmental influences on alcohol use patterns of LGBTQ communities include targeted marketing and advertisements toward these groups and unequal access to quality or LGBTQ-supportive health care. Although attitudes toward homosexuality have become more positive within recent years, studies show that there continue to be some health care providers who would deny health care to LGBTQ individuals based on their sexual minority status and/or hold strong negative opinions about nonheterosexual persons. Even subtle forms of disapproval may be associated with LGBTQ persons struggling with alcohol-related problems to avoid seeking appropriate treatment from a health professional.

PREVENTION AND INTERVENTION

Adolescents and Young Adults

Group, individual, computer- and web-delivered interventions have been found to be efficacious in reducing drinking among several populations. The

most successful brief interventions for adolescent, young adult, and college student alcohol use involve personalized feedback, which incorporates information about one's own drinking patterns in comparison to those of their peers, accurate norms of others' drinking, risk factors for dependence, and negative consequences related to use. Among alcohol programs, these motivational enhancement approaches have reported some of the largest effect sizes. Group, individual, computer-, and web-delivered interventions have also been found to be efficacious in reducing risky sexual behavior. In noncollege samples, the majority of efficacious interventions for risky sexual behavior are delivered in groups and/or in multiple sessions. Despite encouraging results, group and multi-session interventions are more expensive and difficult to implement than web-based interventions. In addition, group and multi-session interventions have limited reach, which supports the need for additional research using efficacious web-based interventions targeting high risk sexual behavior.

Fewer programs have incorporated intervention materials for both alcohol use and sexual behavior simultaneously, although this is an active area of research. Ingersoll and colleagues evaluated an in-person personalized feedback intervention to reduce alcohol-related pregnancy risks among college women, which consisted of personalized feedback for drinking and contraception. At 1 month, women who received the intervention were less likely to engage in heavy drinking, more likely to use effective contraception, and at lower risk for alcohol-related pregnancy. Risk reduction was associated with reductions in the number of drinks per drinking day. Kalichman and colleagues evaluated the efficacy of a single-session group HIV-alcohol risk-reduction intervention among men and women in South Africa. At the 3-month follow-up, findings indicated intervention effects were moderated by alcohol use, such that lighter drinkers demonstrated greater gains from the intervention in comparison to heavier drinkers. Even for effective interventions for risky sexual behavior, intoxication may decrease intervention efficacy. Thus, alcohol interventions shown to reduce drinking may hold particular promise in reducing sexual risk taking.

Sexual Minority Populations

LGBTQ-specific alcohol prevention and intervention services are relatively few despite the recognition that persons in these communities may face additional barriers to reducing or abstaining from alcohol, such as coming out, managing their sexual orientation and gender identity, family and societal stigma and discrimination, coping with health issues that disproportionately affect these communities (e.g. HIV/AIDS), and other comorbid disorders. An approach to assessing alcohol use and alcohol-related problems is to incorporate questions of sexual minority status and behavior into routine medical and psychological screening. Unique considerations when assessing the role of alcohol among LGBTQ patients and clients include their personal experiences during the coming out process, their level of acceptance and comfort with their sexual minority status, social support networks, their relationship to their family of origin, relationship history, past and current sexual behavior, previous treatment experiences, and level of participation in the bar scene.

Alcohol-related treatment models for LGBTQ communities parallel to those of traditional treatment models, with emphases on the disease model of addiction and LGBTQ-specific education and therapy programs. Providing safe spaces to explore the possible association between alcohol use and sexual minority status appears crucial to addressing the needs of LGBTQ clients, which may require the establishment of exclusively LGBTQ treatment programs. In addition, improving the quality of care in programs designed to meet the needs of sexual minorities should include developing LGBTQ-specific outreach programs, sexual minority staff trainings to educate addiction specialists about the unique needs of these communities, and allowances for including partners and families in the treatment process. In addition to these recommendations, it is clear that a greater understanding of the role and impact of alcohol treatment programs tailored for sexual minorities is needed and should be addressed as a future research priority. For example, Velasquez and colleagues found that a tailored individual counseling and peer group intervention focusing on both alcohol use and risky sexual behavior reduced drinking and frequency of alcohol-related risky sexual among MSM.

SUMMARY

The current literature demonstrates the complexity of examining the association between alcohol, sexual risk taking, and HIV/STIs. Numerous variables such as personality traits and characteristics of the sexual partner and context relate to alcohol use and sexual behavior. There is a need for additional event-level research to determine whether alcohol consumption precedes risky sexual behavior and the acquisition of STIs. In addition, event-level research can continue to shed light on which situational variables moderate the alcohol and condom use relationship. Regardless of whether alcohol consumption is a causal risk factor for sexual risk taking and STIs or if alcohol consumption is indirectly related to sexual risk taking and STIs, the literature suggests that individuals who consume

heavier levels of alcohol are at higher risk. Thus, efforts toward developing efficacious interventions aimed to reduce alcohol may be helpful in reducing this risk. Additional randomized controlled trials are needed to definitively demonstrate the causal link between alcohol consumption, sexual risk taking, and HIV/STIs.

SEE ALSO

Adolescent Substance Use: Symptoms and Course, Alcohol and Drug Use in Sexual Minority Youth and Young Adults (lesbian, gay, bisexual, transgender), Alcohol and Sexual Violence, Alcohol's Effects on Sexual Arousal and Sexual Functioning, Binge Drinking, Epidemiology of Addiction, Epidemiology of Adolescent and Young Adult Alcohol, Tobacco, and Drug Use and Misuse in the United States, Families and Addiction, Interpersonal Factors and Addictive Disorders, Minority Groups and Addictions

List of Abbreviations

AIDS	acquired immune deficiency syndrome
CDC	Centers for Disease Control and Prevention
HIV	human immunodeficiency virus
LGBTQ	lesbian, gay, bisexual, transgender, queer, or questioning
MSM	men who have sex with men
STIs	sexually transmitted infections
UAI	unprotected anal intercourse

Further Reading

Cooper, M.L., 2002. Alcohol use and risky sexual behavior among college students and youth: evaluating the evidence. Journal of Studies on Alcohol (Suppl. 14), 101–117.

Cooper, M.L., 2006. Does drinking promote risky sexual behavior? A complex answer to a simple question. Association for Psychological Science 15, 19–23.

Leigh, B.C., Stall, R., 1993. Substance use and risky sexual behavior for exposure to HIV: issues in methodology, interpretation, and prevention. American Psychologist 48, 1035–1045.

Stall, R., Paul, J.P., Greenwood, G., Pollack, L.M., Bein, E., Crosby, G.M., Mills, T.C., Binson, D., Coates, T.J., Catania, J.A., 2001. Alcohol use, drug use and alcohol-related problems among men who have sex with men: the Urban Men's Health Study. Addiction 96, 1589–1601.

Weinhardt, L.S., Carey, M.P., 2000. Does alcohol lead to sexual risk behavior? Findings from event-level research. Annual Review of Sex Research 11, 125–157.

Woolf, S.E., Maisto, S.A., 2009. Alcohol use and risk of HIV infection among men who have sex with men. AIDS and Behavior 13, 757–782.

Relevant Websites

http://www.cdc.gov/ – Centers for disease Control and Prevention (CDC).

http://pubs.niaaa.nih.gov/publications/aa57.htm – National Institute on Alcohol Abuse and Alcoholism.

http://www.effectiveinterventions.org/en/AboutDebi.aspx – The Diffusion of Effective Behavioral Interventions project (DEBI).

http://www.who.int/mediacentre/factsheets/fs349/en/index.html – World Health Organization.

CHAPTER

66

Alcohol Use Disorders

Tian Po Sumantri Oei[*], *Penelope Anne Hasking*[$]

[*]The University of Queensland, St Lucia, QLD, Australia

[$]Monash University, Clayton, VIC, Australia

OUTLINE

DEFINING ALCOHOL USE DISORDERS

Alcohol use has a long tradition in human society, having been used in religious and cultural ceremonies for millennia. Apart from caffeine, alcohol is the most widely used drug in the Western world and is typically associated with celebration, relaxation, and socializing. However, while the majority of people consume alcohol at levels that pose minimal risk to themselves or others, alcohol use can lead to adverse social, personal, and health outcomes, both as a result of acute intoxication and in the long term as a result of chronic use. According to the *Diagnostic and statistical manual of mental disorders 4th edition* (DSM-IV), two patterns of alcohol use are of particular concern: alcohol abuse and alcohol dependence.

Alcohol abuse refers to a recurrent pattern of alcohol use that results in significant impairment in at least one domain of a person's life (e.g. failure to meet commitments at work, relationship problems, legal problems, or financial problems). In addition, a person who abuses alcohol may use alcohol in dangerous situations that place either themselves or others at risk (e.g. driving while intoxicated).

Alcohol dependence refers to a markedly maladaptive pattern of alcohol use. A person who is dependent on alcohol may have difficulty controlling their drinking, feel urges to drink, and continue to consume alcohol despite experiencing adverse consequences (e.g. relationship breakup, losing a job, poor health). In addition, tolerance and withdrawal are defining features of alcohol dependence. Tolerance refers to needing to drink more

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00066-8

alcohol to reach the desired effect. Withdrawal refers to symptoms that appear within several hours to up to 3 days after ceasing to drink. Alcohol withdrawal occurs in several stages. The first signs of withdrawal (including sweating, anxiety, insomnia) may appear within 24 h of ceasing to consume alcohol. Hallucinations and seizures may be present between 1 and 3 days after drinking, while delirium tremens may appear in the later stages of withdrawal. However, it is important to note that the symptoms experienced, their severity, and timing of onset may vary according to the severity of alcohol dependence.

Researchers, clinicians, and policy makers are increasingly expressing concern about binge drinking, particularly among young people. Binge drinking is associated with elevated levels of violence, hospital admissions for alcohol poisoning, traffic accidents, memory loss, and unsafe sexual practices. Unfortunately, there is little consensus on how to define binge drinking. Some define binge drinking as drinking with the primary intention to get drunk, or continual drinking over hours or days. Others attempt to quantify binge drinking as the consumption of five or more drinks within a single drinking session.

In an effort to minimize the health risks associated with alcohol consumption, government agencies have developed guidelines to help educate the public about what constitutes risky drinking. In Australia, the National Health and Medical Research Council recommends that men and women consume no more than two standard drinks each day and encourages several alcohol-free days each week, in order to prevent long-term harm. To prevent acute health risks and injury, it is recommended that men and women consume no more than four standard drinks on a single occasion. In the United Kingdom, the National Health Service advises men to drink fewer than 3–4 units and women to drink fewer than 2–3 units on any given day. In the United States, the National Institute on Alcohol Abuse and Alcoholism recommends up to two drinks per day for men and one drink per day for women and older people. Similar guidelines have been developed throughout the European Union. However, the definition of a standard drink differs across countries. In Australia, a standard drink contains 10 g of ethanol. In the United Kingdom, a single unit of alcohol is approximately 8 g of ethanol, while in the United States, one drink is 14 g of alcohol. Thus, knowledge of the volume of a standard drink as well as the national guidelines is required in order to minimize risky drinking.

PREVALENCE OF ALCOHOL USE DISORDERS

Although rates of alcohol use vary around the world, the World Health Organization estimates that, worldwide, approximately 76.3 million people can be diagnosed with alcohol use disorders. Alcohol abuse is more common than alcohol dependence, and both are more common among males; approximately 5% of the global population may be diagnosed with alcohol dependence at any one time. However, in some countries adolescent girls are drinking at higher rates than their male counterparts, suggesting this gender gap may close in future years.

In most Western countries, alcohol use commonly begins in adolescence. In Australia, 20% of young people aged 14–19 years consume alcohol at least once a week. Rates of abstinence fall sharply between the age brackets 12–15 years and 16–17 years, suggesting the majority of young people begin drinking regularly at about 15 years of age. As early onset of drinking is associated with greater risk of alcohol use disorders later in life, there is a growing concern about the high level of drinking in this age group.

HARMS ASSOCIATED WITH ALCOHOL USE

Alcohol use is responsible for over 2.3 million deaths worldwide, with alcohol-related injuries accounting for one-third of the global burden of disease. Intoxication and alcohol use disorders result in considerable harm as a result of accidents and injuries and have particularly negative consequences in terms of physical and psychological health, as well as social and financial costs.

Accident and Injury

Alcohol plays a major role in road accidents, falls, violence toward others, and self-inflicted violence. In many parts of the world, alcohol is the primary factor implicated in unintended and unwanted sexual intercourse, vandalism, property damage, and accidental death. Accidents and injuries may result from alcohol's effects on coordination, reaction time, and visual focus. However, people who are intoxicated are also more likely to be impulsive, place themselves in high-risk situations, be less averse to taking risks, and have poor judgment and decision making. Together, these physical and cognitive effects of alcohol place intoxicated individuals at significantly increased risk of harm. Importantly, harm resulting from intoxication is rarely limited to the intoxicated individual. The effects on the victims (e.g. other motorists in a car accident, victims of property damage) increase the overall burden of alcohol-related accidents and injuries.

Alcohol use plays a significant role in completed suicide, suicide attempts, and self-harm behavior.

Alcohol abuse among young people who inflict self-harm increases their risk of suicide. The link between alcohol use and self-harm is not clear; however, it has been proposed that alcohol use exacerbates negative affect and minimizes inhibition, which together increase the risk of both fatal and nonfatal self-harm.

Health Effects

Alcohol affects every organ in the body. Chronic alcohol use can have wide-ranging health effects. Alcohol has been linked to a variety of cancers including esophageal cancer, liver cancer, and cancers of the pharynx, larynx, and mouth. There is growing evidence that alcohol use increases the risk of breast cancer in women. As the liver is primarily responsible for the metabolism of alcohol, liver disease, including liver cirrhosis, is a common outcome of alcohol use disorders.

Poor nutrition is common among individuals dependent on alcohol, as essential nutrients usually gained from food are replaced with nutrient-poor, but calorie-rich, alcohol. One nutrient that is often deficient in people with alcohol use disorders is thiamine. Thiamine deficiency may be responsible for many adverse effects including fatigue, mental confusion, and lack of appetite. Poor nutrition may also be partially implicated in alcoholic myopathy. Alcoholic myopathy involves the breakdown of skeletal muscle tissues, leading to tissue loss and pain among people dependent on alcohol.

Alcohol use also has a significant effect on cardiovascular functioning. Hypertension, cardiac arrhythmia, and cardiomyopathy are all associated with increased alcohol consumption. Heavy drinking also increases the risk of hemorrhagic strokes for both men and women.

Drinking alcohol while pregnant places the fetus at significant risk of a number of alcohol-related cognitive and neurological impairments. Fetal alcohol spectrum disorders range from relatively mild cognitive impairment to fetal alcohol syndrome and are 100% preventable. The exact prevalence of fetal alcohol syndrome is difficult to ascertain but is thought to occur in approximately 1% of live births.

Psychological Effects

Alcohol use disorders have been associated with many mental health disorders, and comorbidity is generally considered to be the expectation rather than the exception in those dependent on alcohol. The direction of these relationships is often difficult to ascertain, as mental health issues are known to precede alcohol use, while alcohol use disorders can also result in mental health problems. Arguably the strongest evidence for alcohol use preceding a mental disorder exists for depression. Significant comorbidity between depression and alcohol use disorders exists; however, recent studies suggest that depressive symptoms are alleviated with abstinence, suggesting that at least for some people, alcohol use disorders may increase the risk of depression. There is mixed support regarding a relationship between anxiety disorders and alcohol use disorders. Social anxiety appears to precede alcohol use disorders, with some individuals consuming alcohol for its anxiolytic properties. Withdrawal from alcohol causes anxiety; however, this appears to be a relatively short-term effect.

In addition to causing health problems, lack of thiamine is an underlying cause of Wernicke–Korsakoff syndrome. In Western countries, alcohol use disorders are the primary cause of Wernicke–Korsakoff syndrome, which is characterized by severe memory problems, including both retrograde and anterograde amnesia and confabulation. Wernicke–Korsakoff syndrome is relatively rare, affecting between 1 and 2% of the population.

Consumption of alcohol also disrupts sleep patterns. While intoxication can reduce sleep onset, it can disrupt sleep during the later part of the night. With chronic use of alcohol, the sedative effects appear to diminish, while disruption to the later sleep phases increases. Increased wakefulness and restlessness is related to daytime sleepiness. Withdrawal from alcohol also disrupts sleep patterns, most notably through a rebound of rapid eye movement (REM) sleep. These disruptions to sleep have long-lasting effects, even among individuals who remain abstinent from alcohol after treatment.

Financial Costs

The economic impact of alcohol use disorders is substantial, with the burden on society estimated between 0.45 and 5.44% of gross domestic product. Throughout the European Union, health care costs exceed €22 billion each year. Similarly crime costs (€33 billion), lost productivity (€23 billion), and costs associated with alcohol-related traffic accidents (€10 billion) confer significant economic burden on society. The social costs are also significant and have been estimated to be €270 billion in the European Union, US$185 billion in the United States, and CAN$14.6 billion in Canada.

THEORIES OF ALCOHOL USE DISORDERS

Numerous theories have been proposed to explain the initiation and maintenance of alcohol use as well as the development of alcohol use disorders.

Pharmacology of Alcohol Use

The pharmacology of alcohol use relates to the way in which alcohol acts on the body and the way in which the body responds. The essential premise of a pharmacological view of alcohol use disorders is that people use alcohol because it is inherently reinforcing. Although most people dislike the taste of alcohol initially, social and pharmacological reinforcement encourage continued use of alcohol. From a pharmacological viewpoint, most drugs of abuse, either directly or indirectly, act on dopamine receptors in the pleasure center of the brain. In the case of alcohol, this results in activation of the ventral tegmental area, which produces a release of dopamine in the nucleus accumbens. The release of dopamine in the nucleus accumbens leads to feelings of euphoria and pleasure. Although not as potent as other drugs of abuse, alcohol has an indirect effect on both dopamine and opiate systems. Increased firing of dopamine neurons in the ventral tegmental area is associated with alcohol-induced euphoria. The role of the opiate system is less clear; however, it seems that alcohol use may facilitate the release of natural endorphins, and opioid receptor blockers such as naltrexone reduce the craving for alcohol. However, the primary effect of alcohol is to increase the inhibitory action of γ-aminobutyric acid (GABA), resulting in anxiolytic and sedative effects.

The Role of Genetics in Alcohol Use Disorders

Genetic theories argue that some people are biologically predisposed to alcohol use disorders. The focus is on identifying specific genes, or patterns of genes, that may explain alcohol dependence. The search for a specific gene related to alcohol dependence comes from genetic association studies. Recent studies, in both animals and humans, suggest that the presence of specific dopamine receptors is related to alcohol use disorders.

Genetic Association Studies

Genetic association studies attempt to find a link between a single gene (or group of genes) and a disorder. Researchers attempt to find relationships between individual traits, at least one of which is known to be genetic.

The most consistent genetic association results concern enzymes involved in the metabolism of alcohol. The metabolism of alcohol occurs in several stages. Alcohol is absorbed from the stomach and metabolized by the liver. Here, the enzyme alcohol dehydrogenase (ADH) converts alcohol to acetaldehyde. Acetaldehyde is then converted to acetic acid by the enzyme, aldehyde dehydrogenase (ALDH). Finally, acetic acid is then converted to carbon dioxide and water.

Since the 1970s, researchers have suspected that social drinkers and those dependent on alcohol metabolize alcohol differently and that those who are dependent on alcohol have a genetic fault in the metabolic process. It has been observed that some ethnic groups tend to exhibit an increased flushing response to alcohol. These groups possess an inactive form of the ALDH2 gene (ALDH2*2). Thus, alcohol is metabolized more slowly, resulting in a buildup of ethanol. Those with the inactive form of ALDH2 are thought to be more resistant to dependence due to the aversive flushing reaction that results from the accumulation of acetaldehyde.

Studies have shown a higher frequency of the ALDH2*2 gene in control groups relative to alcohol-dependent samples and that those with the ALDH2*2 gene drink less. These studies have primarily focused on Asian populations, who have a higher frequency of the ALDH2*2 gene. However, it is difficult to separate race effects. ALDH2*2 is not found in Caucasian populations, so this cannot solely explain resistance to alcohol dependence.

Dopamine Receptors

Recently, researchers have investigated genetic differences in dopamine receptors, particularly D2 and D4 receptors. As noted previously, most drugs of abuse, including alcohol, have a direct or indirect effect on dopamine. Thus, differences in dopamine receptors may partially explain susceptibility to alcohol use disorders. The A1 allele of the D2 receptor gene has consistently been associated with alcohol use. This allele is associated with a decrease in postsynaptic dopamine receptors, reducing postsynaptic binding of dopamine. It is theorized that people with this deficiency use or abuse alcohol to compensate and inflate dopamine activation. D2 receptor dysfunction is also associated with a bias toward placing excessive salience on alcohol-related stimuli. As such, alcohol-related cues elicit craving and alcohol-seeking behavior. Similarly DRD2 is associated with a lower self-efficacy for refusing alcohol.

D4 receptors have also been related to alcohol use, but the effects are not as strong or consistent as those noted for D2. Early research suggested that the DRD4 variable numbers of tandem repeat (DRD4 VNTR) polymorphism was associated with urges and craving for alcohol. While some studies have shown a relationship between DRD4 receptor variation and alcohol dependence, other studies have failed to find a relationship. It may be that the effect of DRD4 receptor variation is not a direct one but works in conjunction with other genetic influences, such as those associated with the metabolism of alcohol.

Gene–Environment Interactions

Although the work mentioned above is progressing the search for a genetic factor underlying alcohol use disorders, the role of environmental factors is also important. Family, twin, and adoption studies highlight the complex interplay between genes and environment in the development of alcohol use disorders.

Family Studies

Numerous studies have been conducted to examine the rate of alcohol dependence in children of those who are alcohol dependent. In these studies, the risk of alcohol use disorders in children is three to four times higher if a first-degree relative is dependent on alcohol. These finding are more consistent for males than females. However, alcohol dependence is more prevalent in males, and many studies include only male samples. In addition, although a familial link might suggest a genetic component, family studies cannot rule out other contributing factors such as a shared environment and learned behavior. Thus, familial links are thought to be a combination of genetic and environmental effects.

Twin Studies

Twin studies aim to address the concern that familial links in alcohol dependence may be the effect of a shared environment. If alcohol dependence has a genetic component, then it is more likely that monozygotic (MZ, identical) twins will both develop alcohol use disorders than dizygotic (DZ, fraternal) twins. Overall, twin studies do reveal that the concordance of alcohol dependence is higher in MZ twins than DZ twins, suggesting a role for genetics in alcohol dependence. On average, these studies suggest that between 40 and 60% of variance in drinking can be explained by genetic influences. However, there are several limitations to twin studies that must be considered in assessing the findings, including the assumption that the environment of each twin pair is the same.

In summary, twin studies support the role of both genetic and environmental influences on alcohol dependence.

Adoption Studies

Adoption studies aim to address the criticism that family and twin studies often cannot rule out environmental factors. In adoption studies, children of alcohol-dependent parents who were adopted by nondependent couples are examined. While the children share part of their genetic makeup with a parent who is dependent on alcohol, their environment is not influenced by substance use. These studies have generally found a higher rate of alcohol dependence among children of dependent parents. These rates are again higher for male offspring; however, some support also exists for female offspring.

Behavioral Theories

According to behaviorist theories, alcohol use is viewed as a learned behavior, governed by a set of rules. There are two primary approaches within the behaviorist paradigm: operant conditioning and classical conditioning.

Operant Conditioning

According to operant principles, behavior is learned by associating the consequences of behavior. In other words, if something good happens as a result of engaging in a behavior (reinforcement), we are likely to engage in that behavior again. However, if something bad happens (punishment), we are less likely to engage in that behavior again. Two kinds of reinforcement may increase the likelihood of a behavioral response. Positive reinforcement involves delivering a positive consequence. For example, an individual may gain peer respect by drinking heavily at a party. Negative reinforcement involves the removal of an unpleasant stimulus. For example, the anxiolytic effects of alcohol are a strong negative reinforcer. Both positive and negative reinforcement increase the likelihood of the behavior occurring again.

Operant conditioning has been shown to effectively explain the initiation of substance use. Approval from peers, pharmacological reinforcement, and reduction in aversive moods (e.g. stress) encourage use, and continued use, of alcohol. Further, few people who experiment with alcohol experience punishment as a result of alcohol use (e.g. they are not arrested, do not lose a job, do not suffer negative physical reactions). This combination of reinforcement and a lack of punishment explain why people begin to use alcohol and why they continue to experiment with alcohol. However, one artifact of alcohol use is that these initial effects diminish with repeated use. While alcohol initially results in relaxation, it is likely to increase anxiety over time. As such, an individual can feel the need to drink more to reduce their increased anxiety. Thus, a cycle develops in which a person needs to drink more and more to deal with the anxiety that is actually caused by alcohol.

Classical Conditioning

The classical conditioning, or Pavlovian, approach suggests that people learn behavior by associating previously unrelated events or stimuli. Craving can be a conditioned response to any cue that has been repeatedly paired with alcohol use. In the presence of these

conditioned cues, individuals begin to expect administration of the drug. In the case of alcohol, the body increases the production of the enzymes required to metabolize alcohol; thus alcohol is metabolized more quickly. Consequently an individual needs to consume more alcohol to achieve the desired effect. In other words, tolerance to alcohol can be the result of classical conditioning. If the drug is not administered in the presence of conditioned cues, the body begins to react to the absence of the drug, resulting in withdrawal.

Cognitive Theories

Cognitive theorists argue that behavior cannot be sufficiently explained by stimulus–response relationships proposed by behaviorism. Specifically, our thoughts are believed to mediate the relationship between stimulus and response. Negative or dysfunctional thoughts are the key elements that predict behaviors such as alcohol use. According to Beck's model, either internal (e.g. feeling stressed) or external (e.g. a fight with a friend) events activate the core beliefs that a person holds (e.g. I am unlovable) or beliefs about the effects of a drug (e.g. alcohol helps me relax). These beliefs in turn activate automatic thoughts. Automatic thoughts are fleeting thoughts that everyone has but of which we are often unaware (e.g. a drink will help me relax). These automatic thoughts lead to cravings or urges to use alcohol. This urge to drink leads the individual to construct permissive beliefs (e.g. just one drink won't hurt). Once permission is granted, the person can focus on obtaining alcohol, which then leads to continued use or relapse. This use or relapse can itself become an activating stimulus or lead to modification of other activating stimuli (e.g. change internal mood or external environment).

Social Cognitive Theory

Social cognitive theory (SCT) serves to integrate behavioral and cognitive explanations for human behavior. Central to SCT is the notion that humans do not passively respond to past or current environmental influences but possess the ability to foresee the consequences of our actions. The cognitive outcome expectancies we form allow us to appraise potential consequences of engaging in any given behavior, and thus determine which behaviors we will engage in. Actions we expect will result in positive outcomes are readily adopted, while those that we believe will result in negative outcomes are avoided. This forethought allows us to predict events, set goals, and play an active role in achieving such goals. According to Bandura, self-efficacy is the key factor underlying the human agency central to SCT. Intention to act and a desired outcome are not sufficient to enact behavior; rather individuals must possess self-regulatory mechanisms that allow them to exert control over their behavior. Specifically, regardless of what other factors may affect behavior, unless someone believes they have the ability to perform a behavior or change their circumstances, they will not attempt to do so.

When examining the role of self-efficacy in drinking, researchers have most often studied the role of an individual's belief in their ability to resist drinking (refusal self-efficacy). Refusal self-efficacy is an important predictor of drinking and intention to drink alcohol in children and adolescents. Longitudinal designs have been used to predict teenagers' alcohol and drug use from cognitive and social variables, including refusal self-efficacy. Regardless of whether the participants had experience with alcohol use, refusal self-efficacy was predictive of alcohol use 9 months later. Research by Hasking and Oei has confirmed that in addition to demonstrated salience in clinical and adolescent samples, drinking refusal self-efficacy can discriminate problem and nonproblem drinkers and high- and low-risk drinkers in community samples.

Expectancy Theory

Grounded in SCT, expectancy theory explores the role of anticipated consequences of drinking in the initiation and maintenance of alcohol consumption. Theoretically, individuals who expect positive consequences (positive expectancies) to arise from drinking should consume more, while those who expect negative consequences (negative expectancies) should drink less. The relationship between positive expectancies and drinking behavior has consistently been found in a variety of samples, including adolescents, university students, social drinkers, and dependent drinkers. Jones and McMahon have found negative expectancies to differentiate lone and group drinkers, discriminate satisfaction with current drinking patterns in social drinkers, and predict abstinence and relapse in problem drinkers. Expectations of aggression, risk, affect change, and impaired control have all been related to drinking behavior in social and dependent drinkers.

The Role of Coping in SCT

In the absence of more adaptive coping strategies, a reliance on alcohol to cope predicts drinking behavior. Drinking to cope has been found to be a primary coping strategy predicting alcohol consumption in university students, while a longitudinal study found that drinking to cope was predictive of long-term alcohol consumption and drinking problems in a community sample. Drinking to cope with negative affect has also been

shown to predict DSM-IV alcohol use disorders in a sample of people who did not meet criteria for alcohol abuse at the commencement of the study. More generally, maladaptive or avoidant coping strategies have been related to increased alcohol consumption.

Marlatt and Gordon proposed that after a period of abstinence, being in a high-risk situation without an effective coping strategy results in an individual having a low self-efficacy for their ability to cope with the high-risk situation. Given that an individual has positive outcome expectancies concerning alcohol, this lack of self-efficacy heightens the chance that the individual will engage in drinking. This initial lapse leads to the abstinence violation effect, where individuals question their ability to abstain from alcohol. This process is likely to result in a full relapse to drinking. Hence, Marlatt and Gordon explained how outcome expectancies, self-efficacy, and coping interact to predict relapse of alcohol use disorders.

Hasking and Oei propose that just as coping, outcome expectancies, and self-efficacy are thought to work in concert to determine relapse behavior, they also work in conjunction to govern the decision to drink and the volume of alcohol consumed once this decision has been made. Individual processes such as beliefs about the outcome of drinking, beliefs in the ability to refuse a drink, and coping strategies can explain much of the individual variation in drinking patterns. More importantly, these three constructs have the potential to explain both social and dependent drinking behavior, although the relationships between these variables differ according to drinking patterns.

Intergenerational Transfer

The observation that alcohol use disorders tend to run in families and that cognitions such as outcome expectancies and self-efficacy are the salient predictors of drinking has led Campbell and Oei to propose a cognitive model explaining intergenerational transfer of alcohol use behavior. Campbell and Oei propose that parental behavior (i.e. drinking) influences the formation of the beliefs that their offspring hold about alcohol, which in turn predict alcohol use in the offspring. In addition, the beliefs a parent holds about drinking also influence the cognitions of the child. By observing the drinking behavior of a parent, the child forms expectancies regarding the effects of alcohol consumption. These beliefs, as outlined above, are directly related to the frequency and quantity of alcohol consumption. As such, parents and their offspring share common cognitions regarding alcohol use, and ultimately share similar drinking patterns. This model draws together the findings garnered from family studies as well as those from SCT and provides an explanation for why alcohol use disorders tend to be more common among offspring of parents with alcohol use disorders.

TREATMENTS FOR ALCOHOL USE DISORDERS

The above discussion clearly shows that both genetics and environments are involved in the development, maintenance, and recovery of alcohol use disorders. Alcohol use disorders, in particular alcohol dependence, are complex, chronic, and relapsing disorders. Thus, many treatment approaches ranging from Alcoholics Anonymous(AA) and psychosocial psychotherapy to pharmacotherapy have been developed and implemented. This section discusses the major treatment approaches to alcohol use disorders, in particular alcohol dependence.

Pharmacological Treatment

Many pharmacological agents for the treatment of alcohol abuse and dependence have been identified and tried. They range from baclofen, a GABA-B receptor agonist, to anticonvulsants and mood-stabilizing agents, to disulfiram, acamprosate, and naltrexone. While the list is long, only four such drugs are approved by the Food and Drug Administration (FDA) of the United States of America and many other countries for clinical use in the treatment of alcohol abuse and dependence. These four drugs are acamprosate, disulfiram, oral naltrexone, and ingestible and extended-release naltrexone. The efficacy of these drugs is presented here. Prescription data in the United States indicate that about 1–13% of people suffering from alcohol use disorders received approved pharmacotherapy treatment.

Acamprosate received approval for clinical use by the FDA in 2004. Since then, there have been many randomized controlled trials (RCT) published. A recent review noted that the results of a meta-analysis of 20 RCTs in the United States showed that alcohol-dependent patients receiving 1–3 g of acamprosate had stopped drinking at 6 months follow-up significantly more often than patients who received placebo. However, approximately 36% of patients receiving acamprosate achieved continuous abstinence compared with 24% who received placebo. Thus, the majority (64%) of the patients who had received acamprosate did not stop drinking at all. Other clinical trials completed in Europe show no significant differences in the abstinence rate between patients receiving 2 g (56%) to 3 g (60%) of acamprosate and placebo (54%). A recent systematic review of 24 RCTs found that acamprosate significantly

reduced the risk of drinking compared with placebo, although the effects were modest.

Similarly, patients who received 250 mg of disulfiram drank less frequently than patients who received placebo. There were no significant differences between the disulfiram and placebo groups on the measures of abstinence or time of first drink after treatment. One major point worth noting is that there was a very high noncompliance rate in patients with disulfiram. The rate of nonadherence can be as high as 80%. Another study showed that the nonadherence rate can be improved with the addition of supportive family members in supervision of medication intake. When used in conjunction with family support, disulfiram has been noted to be as efficacious as other pharmacotherapies, at least in the short term.

Oral naltrexone has been used clinically for more than a decade. A clinical dose of 100 mg per day of oral naltrexone could reduce the risk of heavy drinking and also increase the number of days abstinent compared with placebo. However, the effect is modest and not long lasting. In addition, there are numerous side effects of oral naltrexone, including vomiting, diarrhea, and somnolence. Injectable extended-release naltrexone was approved for clinical use to increase the bioavailability of naltrexone and enhance the efficacy of the drug. The literature shows that intramuscular injection of 380 mg of naltrexone had no effect on the number of heavy drinking days but had a modest effect on time to first drink compared with placebo (5 days vs 3 days for the placebo). An advantage of injectable over oral naltrexone is that patients with injectable naltrexone reported minimal side effects. Data from the naltrexone RCTs are somewhat inconsistent, with several large studies finding no significant improvement. However, the majority of studies do report positive treatment outcomes, especially when naltrexone is combined with cognitive behavioral therapy (CBT). A systematic review of 27 RCTs concluded that naltrexone may reduce the risk of drinking in the short term, but evidence was weaker for medium-term and long-term outcomes.

The literature demonstrates that the efficacy of these drugs is generally positive, although the effects are modest. Although disulfiram, acamprosate, and naltrexone are the only FDA-approved drugs for the treatment of alcohol use disorders, other drugs have shown efficacy in treating alcohol use disorders. One of the more promising drugs to emerge is topiramate, an anticonvulsant. A critical review of 26 studies reporting on the use of topiramate to treat a range of substance use problems (12 studies related to alcohol dependence) concluded that topiramate shows efficacy in reducing heavy drinking days, relative to placebo and oral naltrexone. However, only a few RCTs have been conducted, and therefore the research into the efficacy of topiramate continues.

Psychosocial Psychotherapy Treatments

Over the years, a large number of individual psychotherapeutic approaches, ranging from aversive therapy and counseling to CBT, have been used for the treatment of alcohol use disorders. While most individual psychotherapy approaches can claim to have some positive impact on the outcome measures in patients with alcohol use disorders, good solid RCTs confirming positive and sustained long-term treatment outcomes are hard to find. In a large-scale and well-designed multisite RCT, PROJECT MATCH directly compared the efficacy of AA, motivational enhancement therapy (MET), and CBT for the treatment of alcohol-dependent patients. The findings showed that the three therapies had equally significant and sustained improvements in drinking outcomes and that matching specific characteristics of patients to a particular treatment did not improve the drinking outcome.

Similarly many group psychotherapies are commonly used. These can range from self-help delivered by nonprofessionally trained therapists such as AA groups, to manually driven well-structured group CBT, and couples therapy delivered by highly trained professional therapists. Again, group psychotherapeutic approaches, in particular AA, show a high degree of efficacy in drinking outcome, but there is no clear evidence to show that group therapies achieve better drinking outcomes than individual therapies.

Integration of Psychotherapy with Pharmacotherapy

While there is modest evidence to show that pharmacotherapy and psychotherapy alone can have positive outcomes, the literature suggests that combining pharmacology and psychotherapy treatments can produce better treatment outcomes than either pharmacotherapy or psychotherapy alone. The evidence is clearer in showing that the addition of psychotherapy to pharmacotherapy can increase the compliance rate of medication usage and thus can have a positive effect on the treatment outcomes. A recently published RCT demonstrated the efficacy of medical management, with the use of naltrexone and/or behavioral intervention, with the best clinical results observed for the combination of medical management and naltrexone with or without the addition of behavioral intervention. The efficacy of medical management with pharmacotherapy opens the way to general practitioners assisting in the treatment of alcohol dependence.

CONCLUSIONS

There is a consensus that alcohol use disorders are complex, chronic, and relapsing disorders with psychosocial, environmental, and genetic factors contributing to the complexity. However, there is still no consensus on a unifying theory to guide treatment approaches that can provide a sustained and long-term treatment outcome. At present there are too many theories but very little empirical evidence to confirm the validity of such theories. Similarly, there are many treatment approaches, but there is a lack of evidence to show a superior treatment approach that can deliver consistent and sustained treatment outcomes. Thus, improvement and refinement in both theories and treatment are needed. The complex interactions of the psychosocial, environmental, and genetic factors involved in the development and maintenance of alcohol use disorders need further specification, and empirical evidence confirming the validity of theories must be provided. Further empirical evidence for the integration of pharmacotherapy and psychotherapy is needed. It is clear that over the years there has been significant progress in our understanding and treatment of alcohol use disorders. It is also clear that further progress can be made in the future.

SEE ALSO

Binge Drinking, Epidemiology of Addiction, Prenatal Exposure to Alcohol and Illicit Substances, International Data on the Prevalence and Correlates of Comorbid Substance Use and Psychiatric Disorders

List of Abbreviations

AA	Alcoholics Anonymous
ADH	alcohol dehydrogenase
ALDH	aldehyde dehydrogenase
CBT	cognitive behavioral therapy
DSM-IV	*Diagnostic and statistical manual of mental disorders 4th edition*
DZ	dizygotic twins
FDA	Food and Drug Administration
GABA	γ-aminobutyric acid
MZ	monozygotic twins
MET	motivational enhancement therapy
RCT	randomized control trial
REM	rapid eye movement
SCT	social cognitive theory

Further Reading

Bandura, A., 1999. A socio-cognitive analysis of substance abuse: an agentic perspective. Psychological Science 10, 214–217.

Beck, A.T., Wright, F.D., Newman, C.F., Liese, B.S., 1993. Cognitive Therapy of Substance Abuse. Guilford Press, New York.

Campbell, J.M., Oei, T.P.S., 2010a. A cognitive model for the intergenerational transference of alcohol use behavior. Addictive Behaviors 35, 73–83.

Campbell, J.M., Oei, T.P.S., 2010b. The intergenerational transference of alcohol use behavior from parents to offspring: a test of the cognitive model. Addictive Behaviors 35, 714–716.

Garbutt, J.C., 2009. The state of pharmacotherapy for the treatment of alcohol dependence. Journal of Substance Abuse Treatment 36, S15–S23.

Hasking, P.A., Oei, T.P.S., 2002. The differential role of alcohol expectancies, drinking refusal self-efficacy and coping resources in predicting alcohol consumption in community and clinical samples. Addiction Research and Theory 10, 465–494.

Hasking, P.A., Oei, T.P.S., 2004. The complexity of drinking: interactions between the cognitive and behavioral determinants of alcohol consumption. Addiction Research and Theory 12, 469–488.

Hasking, P.A., Oei, T.P.S., 2008. Incorporating coping in an expectancy framework. Current Drug Abuse Reviews 1, 20–35.

Jones, B.T., Corbin, W., Fromme, K., 2001. A review of expectancy theory and alcohol consumption. Addiction 96, 57–72.

Jones, B.T., McMahon, J., 1996. A comparison of positive and negative alcohol expectancies and value and their multiplicative composite as predictors of post-treatment abstinence survivorship. Addiction 91, 89–99.

Marlatt, G.A., Gordon, J.R., 1985. Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors. Guilford Press, New York.

Relevant Websites

http://www.apa.org/helpcenter/alcohol-disorders.aspx – American Psychological Association: Understanding Alcohol Use Disorders and Their Treatment

http://www.niaaa.nih.gov/Pages/default.aspx – National Institute on Alcohol Abuse and Alcoholism

http://www.who.int/topics/alcohol_drinking/en/ – World Health Organization

CHAPTER

67

Heroin Addiction

Andriy V. Samokhvalov[*,**], *Jürgen Rehm*[*,**,$,†]

[*]Centre for Addiction and Mental Health (CAMH), Toronto, ON, Canada [**]University of Toronto, ON, Canada [$]Institut für Klinische Psychologie und Psychotherapie, TU Dresden, Germany [†]Dalla Lana School of Public Health, Toronto, ON, Canada

OUTLINE

PREVALENCE OF HEROIN ADDICTION

Heroin addiction is a major health problem worldwide. According to the data of European Monitoring Centre for Drugs and Drug Addiction 657 t of heroin were produced globally with only 75 t (13%) seized worldwide. Two major forms of heroin historically were manufactured and are offered worldwide nowadays – base form or brown heroin, produced almost exclusively in Afghanistan, and its salt or white heroin, manufactured in Southeast Asia. Heroin purity varies from country to country and correlates significantly with its price and origin, with mean purity of brown heroin in EU being between 15 and 30% and purity of white heroin between 30 and 50%. In USA and Canada heroin purity has significantly increased in the past decade ranging between 40 and 90% in cases of especially expensive shipments.

The most recent and comprehensive systematic review of the data on the prevalence of use and dependence of illicit drugs worldwide was published in 2011 as part of the Comparative Risk Assessment within the Global Burden of Disease Study, and found evidence of opioid use or dependence for 192 out of 229 countries and territories, comprising more than 99% of the world population aged 15–64 years. Past year/point prevalence estimates of opioid dependence ranged from smallest values of around 0.1 per 1000 population in several countries such as Cyprus (0.11), Czech Republic (0.14), India (0.10), and Finland (0.12) to extremely high

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00067-X

values of Iran (8.8) with majority of estimates fluctuating between 0.26 and 1.00 (most European countries). It must be noted that these estimates include, but are not limited to heroin dependence, as many data are based on legal or treatment data.

It has been shown that opioid use has significant impact on social and vocational functioning of the users in high-income countries – patients who cited opioids as their primary drugs when entering treatment have higher rates of unemployment, lower levels of educational attainment and higher levels of psychiatric disorders than those who cited other primary drugs. Almost all opioid users reported initiation of use before the age of 30 and approximately half of them before the age of 20. Estimated time lag between first use of opioids and first contact with drug treatment is approximately 10 years.

PHARMACOLOGICAL AND CLINICAL ASPECTS OF HEROIN ADDICTION

Heroin (morphine diacetate, 3,6-diacetylmorphine) is a semisynthetic opioid produced by acetylation of morphine, natural alkaloid of opium poppies *(Papaver somniferum*). First, it was synthesized by English chemist C.R. Alder Wright in 1874 in London though at that time the substance did not receive any attention. It was independently resynthesized 23 years later by the chemist working for Bayer pharmaceutical company (Aktiengesellschaft Farbenfabriken then), Felix Hoffmann, while trying to synthesize codeine (3-methylmorphine). During the experiment morphine was acetylated that resulted in accidental production of morphine diacetate, which was approximately twice as potent as morphine itself.

Since morphine addiction was a major health care concern at that time and little was known about addictive properties of morphine diacetate, Bayer started its mass production marketing it as nonaddictive and potent morphine substitute under the trade name "Heroin" due to its perceived "heroic" effects on its users. For a little bit more than decade heroin was an over-the-counter medication. Some attempts were made to use it as a treatment for morphine addiction until it was discovered that it is being rapidly deacetylated to monoacetylmorphine and morphine.

After discovery of addictive properties of Heroin its manufacturing and administration was heavily regulated and banned in almost every Western country. Nevertheless, due to simplicity of morphine acetylation it has been efficiently manufactured in illegal laboratories all over the world. It must be noted that heroin is still legally used medically in selected countries.

Pharmacokinetics

Heroin is a prodrug – it is readily deacetylated in human body to 3- and 6-monacetylmorphine and then – to morphine with pharmacokinetic parameters varying to a great extent depending on the route of administration of the drug. The most common route of administration of heroin is intravenous injection that allows for 100% bioavailability of the drug and extremely rapid onset of its action. Due to the presence of acetyl groups diacetylmorphine is more fat soluble than morphine and penetrates blood–brain barrier faster and then is deacetylated to morphine thus ensuring the fastest delivery of morphine to the brain. If administered orally, diacetylmorphine has significantly lower bioavailability and undergoes first-pass metabolism – deacetylation to morphine in the liver. Morphine then, being less fat-soluble substance, penetrates blood–brain barrier slower than heroin would thus resulting in slower delivery of morphine to the brain than in case of intravenous injection. Heroin can also be smoked, insufflated, and in rare cases, used in the form of vaginal or rectal suppositoria.

Pharmacodynamics

Diacetylmorphine and one of its metabolites 3-acetylmorphine have no significant pharmacological effects that are attributable mainly to its fully deacetylated metabolite morphine and to certain extent – to 6-monoacetylmorphine. Both substances have high affinity to and are full agonists of opioid receptors. There are three major kinds of opioid receptors (μ, δ, κ). All three of them are responsible for analgesia, whereas each individual type has several other effects as described below:

μ-opioid receptors were the first receptors identified due to their ability to respond to morphine (μ stands for morphine). In addition to analgesia when activated these receptors cause euphoria – effect sought by heroin recreational users. Also they are responsible for constriction of pupils or miosis, which is one of the hallmarks of opioid intoxication. Two other major effects of μ-opioid receptors activation are constipation (very common in heroin addicts) and respiratory depression (major health risk).

κ-opioid receptors were named after a chemical named ketocyclazocine that binds to them. They also contribute to analgesia, miosis, and sedation, but at the same time they antagonize μ-opioid receptors in respect of mood changes – they cause dysphoria. Though morphine and 6-acetylmorphine bind to both kinds of receptors, effects mediated by μ-opioid receptors predominate resulting in overall euphoric reaction to the drug.

δ-opioid receptors were first identified in murine vas deferens tissues (δ stands for vas deferens). Though they

contribute to analgesia and have certain effects on mood, their role in heroin addiction is minor.

Overall, major effects of heroin are attributable to its active metabolites, morphine, and 6-monoacetylmorphine, and its addictive properties are mediated mostly via μ-opioid receptor agonism.

Heroin Intoxication

Heroin users experience similar symptoms of intoxication every time heroin is being injected, ingested or snorted with some variations depending on the route of administration and heroin experience. The most prominent signs and subjectively perceived symptoms of heroin intoxication are euphoria, relaxation and slight derealization – real stimuli lose their acuity and become numb. These symptoms are most prominent in heroin-naïve users, whereas experienced users tend to need higher doses of heroin in order to achieve the same level of euphoria with repetitive uses as tolerance to heroin builds up very quickly, especially in respect for desirable effects. Historically heroin was used for its analgesic properties, but nowadays it is prescribed for pain control only in selected countries and mostly as a palliative treatment. It doesn't seem likely that someone would start using street heroin for pain management because of its high street value, health risks associated with injection, social and legal risks, tolerance to analgesic effects of heroin and availability of other, cheaper, safer, and legally acceptable options

In addition to the main effects – euphoria and analgesia – heroin use elicits several other symptoms that create a very typical picture of heroin intoxication. They include dryness of mouth and miosis, or pupillary constriction, often described as "pinpoint" pupils of heroin addicts; inhibition of respiratory center leads to lower respiratory rate; decreased gastrointestinal tract motility leads to chronic constipation.

It must be noted that the picture of heroin use varies depending on the route of administration. Intravenous injection is preferred by many users due to pharmacokinetics of intravenous heroin administration described above – rapid onset of action and prominent peak of euphoria in the first moments that together to histamine-mediated vasodilation is described by heroin addicts as heroin flushes. When ingested heroin's effects are not that acute and there is no wave or flush of euphoria. Also due to lower oral bioavailability higher doses are needed in order to achieve similar effect that makes oral heroin use more expensive. Altogether these features of oral heroin use make it less likely to be administered this way. At the same time insufflation, smoking, and rectal suppositoria allow heroin to circumvent its first-pass metabolism in liver and thus are close in perceived effects to intravenous use. Nevertheless, intravenous heroin injection still remains the major route of its administration and is a cause of major adverse effects associated with heroin use.

Individual-Level Harm and Socioeconomic Burden Associated with Heroin Use

Primary set of adverse effects is related to heroin use itself and includes development of withdrawal symptoms after short periods of not using the drug, creation of tolerance and associated risks of overdose and death as well as a number of direct effects of heroin as an opioid agonist as described above in heroin intoxication section.

Besides direct effects of heroin use, it is also associated with a number of risks related to intravenous injections i.e. risk of acquisition of HIV, hepatitis C or any other blood-borne infections. Injections are also associated with a high risk of infection with host's own skin flora such as *Staphylococcus aureus* that is a cause of infectious endocarditis – a hallmark of injection drug use. Another microorganism, *Streptococcus pyogenes*, often causes formation of purulent lesions, abscesses. Chronic injections of heroin lead to deterioration and gradual sclerosis of superficial veins normally used for injections and experienced heroin addicts often have to use veins of lower limbs or groin in order to continue injections. It also complicates the access to the veins for health care professionals, which is extremely important in emergency room settings.

Another common adversity is associated with pharmacologically neutral substances such as talc that are being used for cutting heroin powder. These substances may and often lead to chronic inflammation of blood vessels, vasculitis, that in turn significantly changes the physiological properties of blood vessels and propagates development of atherosclerosis, thrombi formation and thromboembolism, narrowing of the blood vessels section and their increased resistance, etc. Talc and other substances are known for being a cause for pulmonary emboli in heroin use.

The third category of heroin-attributable harm is related to the sequelae of its manufacturing, trafficking, and distribution as well as to the socioeconomic burden related to the costs of treatment, legal proceedings, years of life lost and workforce losses due to disability. All these costs translate into hundreds of billions of dollars spent worldwide on medical services, lawsuits, police enforcement, military operations, disability, welfare payments, etc.

Tolerance to Heroin

Tolerance to various effects of heroin develops with different paces for each effect, i.e. the fastest increase of tolerance is observed with respect to euphoria and

other psychoactive effects. This leads to necessity of constant dose increases, which quickly reach extremely high levels and pleasurable effects of heroin disappear. Heroin analgesic effects require gradual dose increase, but it takes months or years to render heroin absolutely inefficient as analgesic at high doses. Similar pattern of tolerance increase is observed with respiratory center inhibition whereas some, primarily autonomic effects (miosis, constipation) of heroin remain almost unchanged over many years. Such discrepancy in tolerance development inevitably leads to a dangerous clinical scenario when high tolerance to psychoactive effects of heroin leads to extremely high doses of the drug being injected and at the same time tolerance of respiratory centers is not as high and risk of respiratory depression is extremely high.

Heroin Overdose

Overdose is a very common condition among heroin users due to several reasons. One of the major causes of overdose is unevenness of development of tolerance to various effects of heroin described above. Another reason is concomitant use of several substances potentiating psychoactive effects of heroin as well as respiratory depression associated with its use. The most common substances are alcohol, benzodiazepines and barbiturates, antihistamines and sleeping pills (e.g. zopiclone). The third reason is sudden dose increase that can be caused by changing route of administration or process of preparation, switching to a new supplier who provides the user with high purity heroin or banal mistake.

Clinically heroin overdose is characterized with extreme sedation and drowsiness – patients are often described as being "nodding off" as if falling asleep. Pinpoint pupils are classical presentation of heroin overdose as well as any other opioid or opiate. Another set of symptoms is related to respiratory center inhibition – shallow breathing with low respiratory rate (usually below 12) and cyanosis often accompanied with bradycardia.

Overdose Management

Management of heroin overdose is based on identification of the clinical problem and reversal of heroin effects. The first task is to associate sedation with heroin use – presence of a relative, a friend or simply witness of drug use as well as paraphernalia for injection or drug per se can be of great service for quick decision making. Clinical picture of sedation, respiratory depression in association with pinpoint pupils is very patognomic for opioid overdose. Needle marks might be indicative of heroin injection, but there is little clinical importance in identification of particular opioid for overdose management. Possibility of concomitant use of other drugs should be considered and blood toxicology screen should be ordered. Effects of heroin are reversed by intravenous injection of naloxone – potent short-acting opioid receptor antagonist that has higher affinity to opioid receptors than heroin and displaces it from its binding sites. In majority of cases several doses of naloxone are required in order to rouse the patient as well as its repetitive administration in order to achieve stable result due to a very short half-life of naloxone.

Administration of an oral long-acting opioid antagonist such as naltrexone did not become too common due to difficulties with administration of oral medication to semi- or unconscious patient and longer lag time till the onset of action due to absorption. Also, long-acting opioid antagonists put patient in prolonged and subjectively insufferable withdrawal and are feared among drug users.

Heroin Withdrawal

Withdrawal symptoms are generally the opposite of the symptoms of intoxication and are mainly attributable to hyperreactivity to sympathetic nervous system. Normally withdrawal begins several hours after the last dose of heroin, peaks in 2–3 days and resolves mainly within 5–6 days with some symptoms lasting for up to 10 days. Addicts experience extreme anxiety, restlessness, insomnia, fatigue, and strong cravings for heroin. Analgesic effects of previous administration of heroin result in lowered pain threshold and those in withdrawal experience severe muscular and joint pain. Multiple other symptoms are typical for heroin withdrawal – hypersecretions from multiple glands and hypermotility of gastrointestinal tract resulting in sweating, lacrimation, rhinorrhea, diarrhea, and vomiting. Characteristic features are also dilated pupils, chills, piloerection, tachycardia, and hypertension. It must be noted that though heroin withdrawal is subjectively hard to withstand it doesn't lead to any serious medical complications except for cases with somatic comorbidities.

Withdrawal Management/Detoxification

There are several treatment options for heroin withdrawal management often called detoxification or detox. The simplest approach is based on the management of most prominent symptoms of withdrawal preferably in controlled environment – sedatives are used to cope with anxiety, restlessness, and insomnia, often short

course of benzodiazepines is used for this purpose. In order to relieve pain nonsteroidal anti-inflammatory medications are prescribed, nausea and vomiting usually respond well to antiemetics. Diarrhea can be managed with loperamide – opioid that cannot be absorbed from gastrointestinal tract and thus acting locally. Physiotherapy and supportive care are provided.

There are also methods of heroin withdrawal management based on specific medications such as clonidine, methadone, and buprenorphine.

Clonidine Detoxification

Clonidine is an α_2-adrenergic receptor agonist. α_2-adrenergic receptors are located on presynaptic membrane and act as a biofeedback mechanism – each time catecholamines such as epinephrine, norepinephrine or dopamine are released in a synapse they also activate α_2-adrenergic receptors that in turn inhibit further neuromediator release. Thus, clonidine reduces adrenergic tone by stimulating α_2-adrenergic receptors and inhibiting release of norepinephrine, which is the major pathological basis for development of heroin withdrawal.

Usually patients are administered 0.1 mg of clonidine orally with a blood pressure being checked 1 h after that. Since one of the effects of clonidine is hypotension, clonidine is not administered to those whose blood pressure drops below 90/60 after initial dose. If the blood pressure is still above the threshold, clonidine is prescribed in doses of 0.1–0.2 mg orally 3–4 times per day. Duration of treatment is normally limited to several days with potential extension to 8–10 days in total. If doses higher than 0.2 mg four times a day are needed or significant blood pressure drops are observed inpatient treatment is indicated.

Methadone Detoxification

Methadone is a full agonist of opioid receptors with a half-life significantly longer than one of heroin (see Table 67.1: Comparison of opioid agonists). Slow onset of action and longevity of methadone effects almost completely deprive this opioid of addictive properties, though it still has some street value. Due to high potency of the drug extreme precautions are taken when it is prescribed for either detoxification or maintenance. In many countries special license is needed in order to prescribe methadone.

The rationale of methadone detoxification is based on replacement of short-acting highly addictive heroin with the substance of significantly lower abuse potential and less intense withdrawal symptoms due to longer period of elimination allowing for subsequent quick tapering off of it. Due to multiple health risks associated with methadone inpatient detoxification is preferred when methadone is being used. When performed in controlled environment, recommended initial methadone dose is 5–15 mg administered every 6 h for the first 2–3 days. Prior to each administration opioid withdrawal scale is used in order to estimate the severity of withdrawal and necessity of additional dose of methadone. After relatively stable state has been achieved, methadone dose is reduced by 1/3 of previous day's dose and then – tapered off by 5–10 mg day^{-1} unless the patient is to be transferred to methadone maintenance program (see Methadone maintenance).

Outpatient detoxification with methadone is more risky and thus more complicated. Initial doses of methadone are 15–30 mg day^{-1} with subsequent titration – increments of 10–15 mg every 3–5 days. When optimal dose is achieved patients are being tapered off methadone at the rate of 5 mg day^{-1} every 3–14 days. Since there is no direct control over patient's substance use dose titration and tapering should be guided by regular urine toxicology screening. Supportive symptomatic treatment can also be provided.

Buprenorphine Detoxification

Buprenorphine is a newer medication that has been approved for clinical practice in most Western countries with different restrictions. Buprenorphine is also an opioid receptor agonist with a half-life even longer than one of methadone that allows its administration every other day. Unlike methadone it is a partial agonist that contributes to a safer pharmacological profile of buprenorphine with significantly lower risks of sedation and respiratory depression. On the other hand partial agonism of buprenorphine complicates initiation of treatment – when prescribed to someone who still has opioids in their system, it reverses effects of opioids and precipitates withdrawal. In order to prevent precipitated withdrawal buprenorphine should be administered only to those who haven't been using opioids for 24–48 h. In practice this means that these people should be in moderate withdrawal prior to administration of buprenorphine, which in turn will alleviate withdrawal symptoms.

Similarly to methadone buprenorphine detoxification can be performed in either in- or outpatient settings with preference still made to more controlled environment. In inpatient unit buprenorphine is administered to patients who reached moderate withdrawal and quickly titrated to the doses of 16–24 mg day^{-1}. When patients stop experiencing withdrawal symptoms dose of buprenorphine is being quickly tapered off at the rate of 1–4 mg day^{-1} every 1–2 days depending on their condition. In outpatient setting both titration and taper are more spaced in time: initial dose can be up to 8 mg day^{-1} and increased by 2–4 mg day^{-1} every other day until it reaches optimal level. Tapering is usually performed at the rate of 1–4 mg day^{-1} removed every

TABLE 67.1 Comparison of Heroin, Morphine, Methadone, and Buprenorphine

	Heroin	Morphine	Methadone	Buprenorphine
Trade names	Heroin NB: Manufacturing of heroin was mainly discontinued	Kadian Oramorph Avinza Roxanol Kapanol MS Contin	Methadol Dolophine Amidone Heptadon Symoron Physeptone Amydone Phy Methadose	Subutex Suboxone (with naloxone at 4:1) Temgesic Buprenex Norspan Butrans
Routes of administration	Intravenous Inhalation Intramuscular Oral Rectal Intranasal Transmucosal	Intravenous Intramuscular Intrathecal Subcutaneous Oral Inhalation Insufflation Rectal	Oral Intravenous Sublingual Rectal Insufflation	Sublingual Intramuscular Intravenous Transdermal Intranasal
Mechanism of action	Full agonist of opioid receptors	Full agonist of opioid receptors	Full agonist of opioid receptors	Partial agonist of opioid receptors
Active metabolite(s)	6-monoacetylmorphine, Morphine	None	None	Norbuprenorphine
Inactive metabolites	3-monoacetylmorphine Morphine-3-glucuronide Morphine-6-glucuronide	Morphine-3-glucuronide Morphine-6-glucuronide	EDDP	Buprenorphine and norbuprenorphine glucuronides
Bioavailability	≈35–60%	≈25%	≈40–90%	≈35–60%
Protein binding	None for heroin per se	30–40%	85–90%	≈96%
Metabolism	Hepatic	Hepatic	Hepatic	Hepatic
Excretion	90% renal as morphine glucuronides 10% biliary	90% renal 10% biliary	Renal	Biliary and renal
Half-life	≈10 min	2–3 h	12–48 h with great variability	20–70 h
Onset of action	Seconds	Seconds	30 min	30–60 min
Peak of action	10–15 min	20–30 min	3–4 h	1–4 h
Duration of action	4–6 h	2–4 h	24–36 h	Dose-dependent: 8–72 h
Analgesia	Potent, short	Potent, short	Potent, long	Potent, long
Psychoactive effects	Prominent	Prominent	Weak-to-moderate	Weak-to-moderate
Tolerance development	Fast	Fast	Long	Long
Development of withdrawal	Fast	Fast	Slow	Slow
Addictive potential	High	High	Low	Low

several days. Supportive symptomatic treatment can also be provided.

Maintenance/Substitution Treatment

Rationale and methods of maintenance treatment overlap in many aspects with methadone or buprenorphine detoxification though it pursues a different goal – harm reduction versus abstinence in case of opioid-assisted detoxification. Though the same substances are generally used – methadone or buprenorphine – the schedule of titration has a different timeframe; in fact, maintenance treatment may last for infinite time period as long as it prevents harm

associated with illegal activities related to drug-seeking and using behaviors and their sequelae. Also the use of prescribed oral opioid in better controlled environment is intended to prevent risks associated with drug injections and overdose. Opioid maintenance treatment combined with psychosocial interventions was found to be the most effective treatment option for opioid users. In comparison to detoxification methadone or high-dosage buprenorphine treatments show significantly better outcomes in terms of drug use, criminal activities, risk behaviors, HIV transmission, overdoses, and overall mortality as well as better rates of retention in treatment. It is estimated that methadone is used in approximately 75% of maintenance treatment cases and buprenorphine in 20–25% of cases.

Methadone Maintenance

Methadone has been used for opioid maintenance treatment for decades and is well-recognized treatment option in many countries. At the same in some countries (e.g. post-Soviet Union countries) it has never been accepted and is still labeled as narcotic and prohibited.

Due to high potency of methadone and associated risks its use is heavily regulated in most authorities. Unlike short detoxification, maintenance treatment is less aggressive and doesn't require inpatient settings. Normally initial doses are between 15 and 30 mg of methadone administered daily. This dose is not sufficient in most of the cases and is increased slowly at the rate of 5–15 mg added every 3 or more days until it reaches optimal dose. In majority of cases doses of 80–120 mg of methadone are sufficient for maintenance. At the same time small portion of patients who metabolize methadone quickly might require significantly higher doses. Usually during the first weeks or even months of methadone maintenance doses are daily dispensed and taken under supervision of pharmacist. Later when patient's condition is more or less stable and regular urine toxicology do not show any opioids or sedative substances other than prescribed patients are allowed to have take-home doses or "carries" with the maximum of 6 per week so that at least one dose of methadone is observed every week.

In many cases dose of methadone can be partially or completely tapered over the period of months and maintenance therapy resembles extended detoxification. It is considered to be less stressful than detoxification due to longer period of taper and less intense or in some cases no cravings and withdrawal symptoms.

Special attention is paid to potential drug interactions – patients should be educated regarding the side effects of methadone and higher risks of QT-interval prolongation, sedation, and respiratory depression in case of the use of benzodiazepines, alcohol, and other substances with known sedative effects (antihistamines, zopiclone) or those that cause prolongation of QT-interval (some antiarrhythmics, tricyclic antidepressants, etc.).

Buprenorphine Maintenance

Buprenoprhine is considered to have significantly safer pharmacological profile than methadone. It is also relatively new medication and its use is not as rightly regulated as methadone. Due to partial agonism buprenorphine has ceiling effect, i.e. it cannot cause full activation of opioid receptors and hence – dangerous side effects are observed very rarely and in most of the cases are associated with concomitant use of other substances, most often – alcohol and benzodiazepines.

Due to safer profile initiation of treatment and titration of buprenorphine are very fast even in outpatient settings – usually initial dose ranges from 4 to 8 mg day^{-1} with subsequent increases by 2–4 mg every 1–2 days until it reaches optimal level. Normally doses between 12 and 24 mg are enough to control withdrawal symptoms, though in some cases buprenorphine prescribed in doses of up to 32 mg day^{-1} and in some heavy opioid users full control of withdrawal cannot be achieved due to ceiling effect of buprenorphine. Tapering is performed on as needed and as tolerated basis depending on the ultimate goal of treatment – whether it is harm reduction or abstinence. Carries can be provided sooner than in case of methadone, but the same prerequisites should be met – stable control of withdrawal symptoms and no use of illicit substances and drugs with sedative effects as per toxicology screens. Most of the fatalities while on buprenorphine maintenance are associated with the use of alcohol or benzodiazepines. Patient should also be educated about buprenorphine properties and its interactions with other drugs and illicit substances.

Heroin Maintenance

Recently, a program of heroin maintenance for opioid dependence was launched in several European and North American countries that included approximately 2000 opioid users in total. It must be noted that most of the study subjects did not benefit from conventional treatment options. Despite high addictive potential of heroin preliminary results of this project demonstrate significant harm reduction in those involved in heroin-assisted treatment. Moreover, some publications reported better results in patients treated with injectable heroin in comparison to oral methadone. Despite high scientific quality of these findings, heroin-assisted treatment of opioid dependence still seems very limited due to short action of heroin, which in turn dictates frequent dosing; strong psychoactive effects that are the basis of addiction per se; high street value of heroin increasing

the diversion potential of prescribed heroin; use of injectable heroin creates additional risk of development of abscesses and use of other drugs as well as the chances of transmission of blood-borne infections. Another aspect of injectable heroin use is a high risk of overdose and respiratory arrest that makes inpatient treatment settings almost imperative for this kind of opioid dependence treatment. Also, use of heroin as a maintenance treatment masks the use of street heroin, morphine, and codeine in laboratory tests and renders urine sampling virtually ineffective. All these drawbacks of heroin maintenance set significant limits to the use of this modality of opioid dependence treatment and have been proposed for application in specific cohorts of patients, namely, those who were not responsive to other treatment options.

Heroin Detection

Detection of heroin is crucial for medical and legal purposes. In medical settings detection of heroin and its metabolites is diagnostically necessary for confirmation of heroin overdose and heroin dependence. Also regular toxicology screens are required for effective opioid maintenance treatment.

Since heroin is being readily metabolized to monoacetylmorphine and then to morphine testing for diacetylmorphine has not been proven feasible. Instead current laboratory testing methods are aimed to detect heroin metabolites in various bodily fluids that include blood, urine, sweat, saliva, and hair samples testing. While sweat, saliva, and hair samples testing techniques are available, major and most reliable tests for heroin are still based on identification of heroin in blood and urine samples. Blood samples are more reliable and are used during legal procedures and in medical settings when blood samples are taken for multiple purposes or when urine sampling is not possible or complicated e.g. unconscious patient in emergency room. At the same time urine sampling for heroin and other opioids detection is widely used by addictions clinics for confirmation of opioid dependence and for monitoring of treatment progress due to their noninvasive nature and cost-effectiveness. There are two basic methods used for detection of heroin metabolites in urine samples – express qualitative tests based on immunological reactions and more complicated quantitative methods such as gas chromatography–mass spectrometry (GCMS).

Immunological methods are widely used in the form of immunoassays or testing kits for a number of substances. They are cheap, fast, and reliable enough to be used in most of the settings including point-of-care testing of urine samples when samples are processed while the patient is still in the clinic. Since no advanced laboratory equipment is necessary for such testing it becomes more and more popular in smaller clinics. It must also be noted that no detection limit is possible for immunoassays that renders them very sensitive and sometimes leads to false positive results.

GCMS is more complicated and requires expensive laboratory equipment in order to be performed. It renders this technique with several specific properties that in turn define the settings of its usage. First, GCMS is more expensive and can be performed in large laboratory, so no point-of-care testing is possible (some hospital laboratories perform these tests, but even in these settings samples are normally sent to laboratory and it takes at least several hours to retrieve results). Second, due to quantitative nature of testing detection limits are normally set and chances of obtaining false positive results are significantly lower. Third, GCMS produces quantitative results that might be indicative of the dose of the drug used and/or time since exposure.

Normally, individual dose of heroin can be detected 1–2 days after simple use and urine samples remain positive within up to 4–5 days of last use in case of chronic and/or heavy use. It must also be noted that several substances can produce false positive results. These substances can be medical such as codeine, which is available in some over-the-counter or prescribed preparations and is being metabolized into morphine, or even some domestic products such as poppy seeds widely used in baking.

In many cases both immunoassays and GCMS are used simultaneously in order to produce the most reliable results as well as testing for multiple substances is performed. Testing for a spectrum of illicit and medical substances is necessary for several reasons – in case of emergency room testing cumulative effects of several drugs are always considered. Opioid treatment, whether it is detoxification or maintenance treatment, requires confirmation of treatment adherence as determined by presence of methadone or buprenorphine in urine samples and absence of illicit substances. Special attention is paid to substances that can cause sedation such as benzodiazepines, alcohol, barbiturates, antihistamines, antipsychotics, tricyclic antidepressants, etc.

One of the major drawbacks of urine sampling is a risk of urine samples being tampered or switched by patients in order to evade legal consequences or continue use of medical facilities without being penalized (i.e. more structured and restrictive settings of provided maintenance treatment services, loss of driver's license or parental rights) for continued use of heroin despite treatment agreement and goals. Since both methadone and buprenorphine have street value another goal of urine samples tampering might be to continue receiving prescribed opioid for subsequent selling it on black market while using only portion of it for own needs or not using it at all. The most effective

way of prevention of such activities is incorporated in essentials of any maintenance schedule – as at least one dose a week must be taken under supervision of pharmacist users who only use part of methadone or buprenorphine prescribed them on daily basis tend to become extremely sedated when taking full dose and thus are easily identified.

The most common way of urine samples' tampering is providing urine samples of another methadone or buprenorphine user who doesn't use illicit drugs. It can be prevented by several safety checks. The most common of them is supervised urine collection, sometimes performed randomly. Another way to check if urine sample belongs to the patient is to check its temperature – in most of the cases tampered urine samples have temperature significantly lower than freshly urinated ones. Since urine toxicology testing often includes multiple substances, presence of medications not prescribed to the patient is an excellent indicator of urine being taken from another person.

Some users are also trying to dilute their samples with tap water in order to prevent identification of illicit drugs in urine by GCMS. This tampering technique is controlled for by including urine creatinine levels in laboratory reports that when extremely low are indicative of urine dilution. Also, urine dilution does not prevent detection of many substances by immunoassays that are extremely sensitive. Some clinics also use special markers or coloring agents added to the water in their washrooms that render sample dilution virtually impossible.

POLICY

Since the very discovery of addictive properties of heroin, it became a subject for multiple polices aimed to regulate its use and restrict the spread and scope of heroin addiction both in separate countries and worldwide.

Legal Prohibition

Heroin production, distribution, and prescription became illegal in many countries decades ago. It is hard to say how prevalent would be heroin use and dependence without this prohibition, though it is clear that on the one hand it would be significantly more prevalent, but on the other hand it is still unacceptably high. Empirical research has pointed out, that further increases in enforcement and incarceration in many countries would yield very little or no added benefit.

Worldwide Incentives to Reduce Production of Heroin

Several wealthy countries, must notably the USA, have invested billions of dollars in efforts to reduce cultivation of *Papaver somniferum* in poor countries. Unfortunately this did not lead to any significant changes of heroin production in a long run. Inquiry in the US-sponsored programs of agricultural development in Afghanistan showed not only that these programs failed to reduce drug production, but moreover that heroin production has increased the most in those areas receiving such aid. The evidence on mass defoliation campaigns is also pessimistic – sustained and intense campaigns of crop spraying in Colombia and Mexico had little effect on drug production in these countries.

Harm Reduction

Several incentives were proposed in order to reduce heroin-associated harm. Majority of them aimed to reduce the harm associated with injections – needle exchange, providing drug addicts with free needles, needle disposal, etc. These incentives resulted in significant decrease of HIV and hepatitis C infections rate in selected communities, but at the same time left a wide range of medical, social, and legal problems associated with heroin use unsolved.

Treatment for Heroin Dependence

As it was shown above several treatment options were proposed and widely used for management of heroin dependence that include both abstinence-based approach (detoxification) and maintenance treatment options. These strategies imply either complete elimination of heroin use or substitution of heroin, highly addictive short-acting illegal substance that requires higher and higher doses to be injected, – with long-acting oral prescription opioid that is less addictive and tolerance to which doesn't build up as quickly as to methadone or buprenorphine. Both medications have significant advantages in comparison to heroin – they are cheaper or free to the patients in most countries, oral administration eliminates the risks associated with intravenous drug use, they have longer half-life so there is no blood concentration "roller-coaster" and hence these medications give no or very little positive reinforcement and at the same time provide patients with steady levels of opioid effectively preventing development of withdrawal and eliminating the necessity of drug-seeking activities.

School, Family, and Community Prevention Programmes

Multiple attempts were made to prevent the use of heroin, especially its injection, based on educational campaigns targeting certain communities or social

groups, especially children. Prevention programmes that are delivered to younger people before they initiate drug use and targeting a wide spectrum of mental and behavioral problems as well as drug use have emerged over the last 20 years and seem to be the most promising of this kind of intervention. In contrast, purely didactic prevention programmes delivered through mass media, in the community or in the classroom have no evidence of effectiveness. Overall, the impact of individual-level prevention is limited.

SUMMARY

Heroin addiction is a major health problem globally. Though during past two decades it was partially substituted by misuse of prescription opioids in high-income countries, especially in North America, and multiple harm reduction programs were initiated in many countries worldwide, size and burden of heroin addiction remains unacceptably high. Heroin and heroin addiction have been well studied on multiple levels starting with chemical analysis of the substance, its pharmacological properties and effects on human organism to the impacts of heroin use on economy, health services usage and society in general. Methods of heroin intoxication, overdose, withdrawal, and dependence treatment are well established in many Western countries as well as in a number of developing countries. Various forms of policies and programs were introduced with some degree of success in reducing the harm associated with heroin.

At the same time millions of people worldwide begin and continue using heroin, become addicted to it and are involved in various illegal activities related to heroin addiction, significant portion of patients involved in heroin addiction treatment programs either relapse or do not benefit from these programs at all. All these factors point to the necessity to further elaborate effective anti-heroin policies as well as to improve heroin addiction treatment to better meet the needs of heroin addicts.

SEE ALSO

Prescription and Over-the-Counter Medications, International Perspectives on Addiction, Epidemiology of Addiction, Historical Understandings of Addiction, Medical Toxicology of Drugs of Abuse

Glossary

Buprenorphine partial agonist of opioid receptors used for maintenance treatment of heroin addiction.

Clonidine α_2-receptor agonist used for alleviation of heroin withdrawal.

First-pass metabolism metabolism of the drug that happens only in case of its oral administration when the drug is being drained to liver and being partially metabolized by liver prior to reaching systemic circulation. Some drugs, e.g. buprenorphine, have extensive first-pass metabolism that renders them virtually ineffective when prescribed orally.

GCMS gas chromatography–mass spectrometry

Maintenance therapy synonym of substitution therapy.

Methadone full agonist of opioid receptors used for maintenance treatment of heroin addiction.

Naloxone short-acting opioid receptor antagonist used for reversal of excessive activation of opioid receptors in heroin overdose, primarily sedation and respiratory depression,

Opiates natural alkaloids of opium or its semisynthetic derivatives (also called semisynthetic opioids). This category includes morphine, codeine, thebaine, and papaverine.

Opioids synthetic opioid receptor agonists. In many cases this term is used more broadly for semisynthetic opioid receptor agonists or interchangeably with "opiates."

Substitution therapy therapeutic replacement of illegal substance with less harmful prescribed medication, e.g. substitution of heroin with buprenorphine or methadone.

Withdrawal management, or detoxification complex of therapeutic measures aimed at treatment of withdrawal syndrome and ultimately achieving abstinence.

Further Reading

Amato, L., Davoli, M., Perucci, C., Ferri, M., Faggiano, F., Mattick, R., 2005. An overview of systematic reviews of the effectiveness of opiate maintenance therapies: available evidence to inform clinical practice and research. Journal of Substance Abuse Treatment 28, 321–329.

Amato, L., Davoli, M., Minozzi, S., Ali, R., Ferri, M., 2005. Methadone at tapered doses for the management of opioid withdrawal. Cochrane Database of Systematic Reviews 3 CD003409.

Babor, T.F., Caulkins, J.P., Edwards, G., Fischer, B., Foxcroft, D.R., Humphreys, K., et al., 2010. Drug Policy and the Public Good. Oxford University Press, Oxford.

Degenhardt, L., Bucello, C., Calabria, B., Nelson, P., Roberts, A., Hall, W., et al., 2011. What data are available on the extent of illicit drug use and dependence globally? Results of four systematic reviews. Drug and Alcohol Dependence 117, 101.

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA). Annual Report on the State of the Drugs Problem in Europe. Lisbon, 2010. Accessible online: http://www.emcdda.europa.eu/publications/annual-report/2010.

Gowing, L., Ali, R., White, J.M., 2009. Buprenorphine for the management of opioid withdrawal. Cochrane Database of Systematic Reviews 3 CD002025.

Gowing, L., Farrell, M., Ali, R., White, J.M., 2009. Alpha2-adrenergic agonists for the management of opioid withdrawal. Cochrane Database of Systematic Reviews 2 CD002024.

Jones, L., Pickering, L., Sumnall, H., McVeigh, J., Bellis, M.A., 2010. Optimal provision of needle and syringe programmes for injecting drug users: a systematic review. International Journal of Drug Policy 21, 335–342.

Mattick, R.P., Breen, C., Kimber, J., Davoli, M., 2009. Methadone maintenance therapy versus no opioid replacement therapy for opioid dependence. Cochrane database of systematic reviews 3 CD002209.

NIDA., 2005. Heroin: Abuse and Addiction. NIDA, Rockville, MD.

Nutt, D.J., King, L.A., Phillips, L.D., Independent Scientific Committee on Drugs, 2010. Drug harms in the UK: a multicriteria decision analysis. Lancet 376, 1558–1565.

Oviedo-Joekes, E., Brissette, S., Marsh, D.C., Lauzon, P., Guh, D., Anis, A., et al., 2009. Diacetylmorphine versus methadone for the treatment of opioid addiction. New England Journal of Medicine 361, 777–786.

Rehm, J., Gschwend, P., Steffen, T., Gutzwiller, F., Dobler-Mikola, A., Uchtenhagen, A., 2001. Feasibility, safety, and efficacy of injectable heroin prescription for refractory opioid addicts: a follow-up study. Lancet 358, 1417–1420.

Treatment improvement protocols, TIP 40: clinical guidelines for the use of buprenorphine in the treatment of opioid addiction. Accessible online: http://www.ncbi.nlm.nih.gov/books/NBK14901/.

The College of Physicians and Surgeons of Ontario. Methadone maintenance treatment program standards and clinical guidelines, fourth ed., 2011. Accessible online: http://www.cpso.on.ca/uploadedFiles/policies/guidelines/methadone/Meth%20Guidelines%20_Oct07.pdf.

CHAPTER

68

Cocaine Addiction

Robert Malcolm, Kelly S. Barth, Lynn M. Veatch

Medical University of South Carolina, Charleston, SC, USA

OUTLINE

INTRODUCTION

The coca plant, *Erythroxylum coca*, and related species have grown naturally in the northern mountains of South America for thousands of years. The chewing of coca leaves with alkaline substances such as lime powder is an ancient practice of pre-Incan peoples. The improvement in fatigue, reduction of hunger, augmentation of cognition, and enhancement of mood were compelling assets of the plant for indigenous peoples. Processing the leaves into cocaine hydrochloride or alkaline cocaine (crack) and distributing it in modern society highlighted the compelling liabilities of the drug. Stimuli salience, reward, learning, memory, cognitive focus, and emotion are altered by repeated cocaine use. Using to get high is replaced by intense craving and compulsive drug seeking/using, defying legal, social, and medical adverse consequences. At the core of this progression is the intense euphoric and hedonic joy produced initially by cocaine, quickly followed by drug hunger. Repeated seeking and use are based on cocaine's ability to influence and subvert fundamental brain reward systems by redirecting dopaminergic and glutaminergic circuitry. The journey from first exposure to cocaine addiction is a complex one involving individual vulnerabilities as well as cultural, economic, legal, and demographic modifiers.

BOTANY

Cocaine is a complex pharmacologic agent. It has the combined properties of a systemic stimulant, a local anesthetic, and a powerful vasoconstrictor. Cocaine is derived from the leaves of the coca plant, which is indigenous to western mountainous South America. It is farmed at higher altitudes of the Andes and is part of the family Erythroxylaceae. The best-known plant in this genus is *E. coca*, but multiple species and subspecies exist throughout the Andes Mountains. Indigenous peoples of Peru, Colombia, Ecuador, Venezuela, Bolivia, and northwestern Argentina have cultivated, sold, and consumed the plant since prehistoric times. The plant resembles a tall bush, growing to a height of 2–3 m. The leaves are green, thin, and tapered at both ends. In South America, the seeds are generally sown from

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00068-1

December to January in small shaded plots of ground, and then later transplanted to a sunnier location to complete the growing cycle. Plants may be harvested for many years, but only the new growth leaves are harvested. Harvest seasons are normally three times a year: March, June, and November. Leaves are spread in thin layers and dried in the sun, then packed in sacks for transport. The leaves must be kept dry to preserve the quality of the cocaine. Depending on the species and the location grown, the leaf contains about 0.3–1.5% cocaine. Coca plant cultivation has been tried in Java, Taiwan, Africa, India, Ceylon, and Australia.

HISTORY OF COCA AND COCAINE

Traces of coca have been found in Andean mummies dating back to 3000 BC, and other investigators have asserted that traces of coca can be found back as far as 8000 BC. Many archaeological sites reveal widespread coca cultivation and use by the sixth century AD. Coca leaves were routinely buried with the dead. Excavations of pottery and statues illustrate cheeks bulging with the chew of coca leaf/balls as well as paraphernalia, such as spatulas for adding alkali to coca chew, often made from gold or silver. After the Spanish conquistadores came to South America and broke the Inca Empire, coca leaves became a literal cash crop, being the currency exchanged for goods and services. Coca leaf chewing with lime became common among laboring classes of Incas and other indigenous peoples. The Catholic Church in South America banned coca leaf chewing in the early 1500s, but the decision was reversed when the indigenous peoples, enslaved to work in the mines and other hard laboring tasks, were found to work more and ate less when chewing coca leaves. Philip II of Spain issued a proclamation allowing for coca leaf use among indigenous peoples and began to tax the coca leaf. Medical writings of the physician Nicolas Bautista Monardes, published in 1574, detailed the pharmacologic and behavioral actions of coca leaves. Monardes was the first to describe the mixture of a white alkaline powder made from seashell ash with coca leaves rolled into a ball and chewed in this fashion to enhance absorption of the cocaine alkaloid. Discovery of the anesthetic effects of cocaine was attributed to the Jesuit priest Benarbé Cobo in 1653, when he described chewing coca leaves to stop his toothache.

Despite these early writings being circulated in Europe, the chewing of coca leaves with alkali powder never spread to Europe like the smoking of tobacco. Whereas dried tobacco leaf is relatively stable over ocean voyages, coca leaves are not, and their pharmacologic effects were probably greatly degraded by the time they reached Europe. Undoubtedly, the route of administration of nicotine, primarily by smoking, also contributed to its spread and nicotine addiction. The innovation of smoking cocaine took about 400 years.

In 1582, a European account detailed the morbidity and mortality among indigenous peoples using coca leaves on a regular basis. Although some of the mortality was likely due to comorbid infectious diseases, poor diet, and harsh working conditions, this is an early record of the negative consequences of coca leaf use.

In 1859, Albert Niemann isolated the major active alkaloid in coca leaves and named it cocaine. He noted its local anesthetic properties. Sigmund Freud published an influential review paper on cocaine in 1884, *Uber Coca*. Freud at first supported the pharmacologic value of cocaine, using it to treat opioid dependence. Later, after several adverse reactions, including a patient's psychosis, he abandoned its clinical use. Freud had discussions with his friend Karl Koller, an ophthalmologist, about the anesthetic properties of cocaine. Koller is given credit for developing the local anesthetic properties of cocaine by operating on a patient with glaucoma and publishing the outcome. Use of injectable cocaine quickly spread to dental and otolaryngologic applications, and cocaine is still used today for local and regional nerve block anesthesia. From the 1860s to 1900, cocaine was used in medicinal tonics and recreational beverages. In Europe, Vin Mariana, an elixir made from red wine, coca leaves, and possibly refined cocaine, was popular, but frequently led to abuse and addiction. In the United States, John Pemberton created Coca-Cola, which contained coca products until 1903. In 1906, the Pure Food and Drug Act ordered the removal of opioids and coca products from over-the-counter medications in the United States.

EPIDEMIOLOGY OF COCAINE USE

Coca leaf ingredients and the alkaloid of cocaine itself were frequently used in patent medicine products and popular drinks from the 1880s until the early 1900s. This was the first cocaine epidemic. Exposure to cocaine and coca leaf products was extensive until the Pure Food and Drug Act of 1903 and the Harrison Narcotic Act of 1914. Cocaine use remained at a low level until the 1960s, when nasally used cocaine hydrochloride became popular among upper income classes. Organized drug cartels, seeing a market for the drug, began importing large quantities of cocaine hydrochloride from South America, and the use of nasally insufflated cocaine peaked in the mid-1980s. Reports of cooking cocaine with alkaline substances (i.e. baking soda) began appearing in the early 1970s, and by the mid- to late 1980s, smoking of rock, more commonly known as crack

cocaine, became widespread. Cocaine base (crack) is easy to transport, is highly potent, and may be smoked to give effect on the brain in a few seconds. The cost of crack was lower than nasally insufflated cocaine hydrochloride and could therefore be sold to lower socioeconomic groups. Laws passed during this time were more severe against crack users compared with laws against nasal use of cocaine hydrochloride. Only recently are laws becoming more equable. For reasons that are not clear, all forms of cocaine use have declined over the last decade. Prevention programs, public media campaigns, and law enforcement all prefer to take credit for these trends. By 2005, fewer high school students had tried cocaine than in the previous decade. About 34 million individuals had used cocaine at some point in their life, almost six million individuals had used cocaine in the past year, and 1.5 million individuals used some form of cocaine on a regular basis, largely crack cocaine.

Use among males is usually two or three to one compared with females. The rate of unemployed individuals using cocaine is double that of employed individuals. Cocaine use by males generally begins in the late teen years to the early 20s; in females it begins about 5 years later. Cocaine use in females is said to have a telescoping effect in that use begins later but more devastating psychosocial effects from cocaine occur sooner. The Drug Abuse Warning Network reported about 78 emergency department visits out of 1000 were related to cocaine use. For comparison, about 7 per 100 000 were said to be methamphetamine related. These figures are probably low estimates in that individuals who develop chest pain involving cardiovascular or pulmonary symptoms, or for that matter adverse psychological symptoms, will often go to an emergency room and will not admit that cocaine is involved unless a urine drug screen is done. An epidemiological curiosity is that generally cocaine use is higher east of the Mississippi, and methamphetamine use is higher west of the Mississippi. Chicago, Philadelphia, Baltimore, and Miami are known as cocaine cities. Methamphetamine cities include San Francisco, Seattle, Los Angeles, and Phoenix. Possibly, these patterns of cocaine and methamphetamine use are related to organized crime distribution issues.

PHARMACOLOGY OF COCAINE

Cocaine is an ester of benzoic acid and methylecgonine. Ecgonine is an amino acid base. Both topical application and local injection of cocaine solution blocks the generation and conduction of nerve impulses. The site of action is the cell membrane where cocaine prevents conduction by reducing membrane permeability to sodium. This blockade of sodium channels increases the threshold for electrical excitability, which reduces the probability of nerve impulse transmission. In addition, cocaine inhibits the reuptake of norepinephrine conferring vasoconstrictive effects as well. The significance is that other local anesthetics do not have this property of vasoconstriction and generally must be given with a separate vasoconstrictive agent. Cocaine is used primarily as a topical anesthetic in dental, ear, nose, and throat procedures. The combined vasoconstrictive and local anesthetic properties make cocaine an ideal agent providing anesthesia and shrinking swollen mucosa.

Cocaine metabolism is primarily hydrolysis of ester groups. Esterases are found in blood, and benzoylecgnine (the major metabolite of cocaine found in urine) is detectable up to 2–5 days after binge use. Cocaine itself can only be measured in the serum from minutes to 2–3 h after use. Rarely, heavy users may have cocaine metabolites for up to 10 days after use. A minor pathway of cocaine metabolism is through a CYP3A4 mechanism with degradation to norcocaine.

The absorption of cocaine varies greatly with its route of administration. Oral cocaine (or coca leaves) is absorbed into the central nervous system in 30 min to 1 h. Nasal cocaine is absorbed in 2–3 min. Smoked cocaine and intravenous cocaine reach the brain in 10–30 s. With smoked cocaine, craving and new drug seeking may occur as quickly as 5–10 min. Positron emission tomography studies of cocaine abusers using cocaine labeled with carbon 11 demonstrated that the time course of the subjective euphoria of cocaine paralleled the uptake of the drug in the corpus striatum (nucleus accumbens and related structures).

Clinical Pathology and Course of Cocaine Use

Cocaine intoxication is described as a high or rush feeling, and other effects include heightened euphoria, rapid pulse, increased blood pressure (more systolic than diastolic), and increased motor behavior and talking. Greater alertness, greater anger, increased anxiety, greater gregariousness, increased vigilance, poor judgment, and labile mood also occur. In heavier users, stereotyped repetitive behaviors (i.e. rubbing one's nose, rubbing one's forehead) or more complex changes of behavior occur. Repetitive cocaine users will often say that their first high was the best, and no high approaches the first one, hence the phrase, "chasing the [first] high." A large amount of cocaine leading to overdose, on any one acute use, can be associated with medical complications, which can include death from cardiac arrhythmias, seizures, hyperthermia, dehydration, and localized vasoconstriction leading to myocardial infarction or stroke. Some paranoia is common with repeated

cocaine use. In some users, this may increase to a paranoid psychosis lasting for days or weeks.

Recurrent users of cocaine have withdrawal symptoms, but they are not as well defined as can be seen with alcohol, sedative-hypnotic, or opioid dependence. Following the end of a cocaine binge, which may last for hours to days, individuals have dysphoric mood, decreased motor activity, social withdrawal, exhaustion, bradycardia, and lowered blood pressure. Symptoms may persist for a few days after cessation. Complaints of dysphoric mood, lack of pleasure (anhedonia) from biological and social rewards may persist for weeks to months to years. Drug craving, increased appetite, or unpleasant dreams may occur in the same interval. Some cocaine users report immediate spontaneous craving and drug seeking after use. Others deny any conscious craving but also quickly begin seeking and use. Another segment of users report no use unless stressed (an argument, work demands, financial or legal troubles) or very specific cues occur (cash, a dealer, another user, a drink of alcohol, a street, or a building). The list of stressors and cue triggers can appear infinite to family and therapists.

Physical examination findings in cocaine users are not that common or diagnostic. Acute users will have tachycardia, increased blood pressure, and motor agitation. Heavy intranasal users will occasionally have an eroded nasal septum and/or frequent nosebleeds. Injection users have the usual signs of frequent needle use on arms, legs, or neck, and concurrent infections with hepatitis B or C and human immunodeficiency virus occur. Crack smokers will occasionally have burns on the hands and lips. Urine drug screens have some benefit in detecting cocaine use, but again they are limited by the rapid metabolism and excretion of benzoylecognine and the requirement to test individuals three or more times weekly. Episodic bingers may go for years with occasional use, without detection by employers, probation and parole agencies, and the military.

The clinical course of cocaine addiction is extremely variable. Inexperienced younger users are more likely to be victims of the complications of overdose and to develop medical, psychiatric, and legal difficulties. Older chronic users, who have survived the entanglements, may present for treatment in relatively good health after more than two decades of use. It is common for that latter group to be underemployed, estranged from families and significant others, and to have significant comorbidities with alcohol and marijuana. The course of cocaine use from initial use to chronic use is extremely variable. The authors know of individuals who, on the very first use of intravenous cocaine, suffered cerebral hemorrhage and died. Other patients come to mind who used cocaine on one single occasion, probably suffered a fatal arrhythmia, and died shortly thereafter. Other individuals have moved gradually from snorting cocaine hydrochloride with a rather stable pattern of recreational social use, predominantly at weekends over a number of years, to intravenous use or, in more recent years, smoking crack cocaine with devastating social and medical consequences. One individual who comes to mind lived in a wealthy resort area, owned several regional businesses, had a house on the beach and one in the mountains, and several horses boarded in stables. After switching from many years of nasal cocaine use to smoking crack cocaine, he lost both of his houses and his riding stables in less than a year. Spending money on cocaine for himself and others, poor business decisions, and the alienation of key people in his social network led to his rapid socioeconomic decline.

Cardiovascular complications are common with acute cocaine use. Chest pain and tachycardia are frequent reasons why individuals go to emergency departments. Myocardial infarctions are relatively uncommon but can occur in people with preexisting heart disease who are often cigarette smokers; myocardial infarction may also occur in individuals with normal coronary arteries. Patients developing acute ventricular arrhythmias from cocaine probably do not live to make it to emergency rooms. Cocaine use has been associated with dissection and rupture of the abdominal aorta. Frequently in cigarette smokers, peripheral vascular disease has been associated with chronic cocaine use. Sometimes this is preceded by accidental injection of cocaine into an artery.

Smoking crack cocaine can cause acute constriction of the bronchi and may lead to wheezing and attacks indistinguishable from asthma. Increased blood pressure in the pulmonary arteries has been reported in intravenous cocaine users. Nasal users of cocaine have been reported to have nasal septal erosions, ulcers, and perforations. Oronasal fistulas have been reported, and perforations of the hard palate have also been seen. Worsening of kidney function has been noted occasionally in chronic cocaine users, especially in individuals with high blood pressure who are crack smokers.

Cocaine can produce numerous neurological and muscular complications. All types of strokes have been seen with cocaine use. Ischemic stroke leading to clot formation has been reported. Intracerebral bleeds have been reported, and subarachnoid hemorrhages occur in some cocaine users. A transient ischemic attack or a stroke in an otherwise young healthy individual without risk factors should suggest cocaine or other stimulants as part of the differential diagnoses. Heat stroke and other forms of hyperthermic injuries can occur in cocaine users. This is usually in the context of high environmental temperatures, hyperactivity, lack of fluid intake, and possibly impaired heat loss from lack of sweating.

In vulnerable individuals, cocaine use can lead to a variety of psychiatric syndromes including generalized anxiety, panic attacks, and psychosis. Agitation and paranoia are frequent in cocaine users and may persist for hours to days after binge use.

In summary, the clinical course of cocaine dependence varies greatly depending on individuals' vulnerabilities, socioeconomic factors, co-occurrence of psychiatric and medical comorbidities, as well as issues of polysubstance use.

Antecedents of Cocaine Dependence

There are social, psychological, and biological factors operating simultaneously to affect the initiation, perpetuation, and course of cocaine dependence. Some of the social factors that lead to the initial use of cocaine have been hinted at in the section on epidemiology. Humans are social animals, and it is usually in the context of a social setting, peer pressure, modeling of other individuals' behaviors, desire for greater acceptance, and exploratory and experimental behaviors that cocaine use is initiated. Cocaine use usually occurs in the setting of some other previous drug experience such as alcohol, marijuana, or prescription medicine use. For clinicians, etiologies of cocaine abuse and dependence are only valuable insofar as they may point to meaningful treatment interventions that lead to effective treatment outcomes for individuals with addiction to cocaine. The complexities of genetics, public policy, local social norms, geopolitical events, and economic conditions play a role in the cause and perpetuation of cocaine addiction. These subjects and the neurobiological foundations of cocaine dependence are examined in detail in other chapters in this volume.

Koob and Kalivas refer to the social drug use as regulated relapse and a more serious second phase of almost automatic drug use as compulsive relapse. For a clinician, the first might be defined as regulated drug seeking and use and the second as compulsive drug seeking and use; relapse is reserved for an extended period of abstinence with or without treatment that is followed by drug use again.

There are some important distinctions about the release of dopamine in the nucleus accumbens by cocaine compared with dopamine release by physiologic rewards such as food, water, or sex. In animal research, release of dopamine with cocaine achieves a greater amplitude and duration of dopamine release than biological rewards. Tolerance rapidly develops to the release of dopamine by food, sex, or water reinforcers. Although tolerance develops to cocaine-promoted release of dopamine, this generally can be overcome by a higher dose of the drug, enabling the same enhanced level of dopamine release. Satiation to biologically motivated reinforcers such as food, sex, or water generally occurs quickly, although some human pathological conditions, such as binge-eating disorder and bulimia, are exceptions in which these conditions mimic drug addiction. Each time cocaine is used, expanded learning occurs, strengthening the associations between the environment and the drug. This leads to a myriad of cues that will later trigger the addicted individual to begin drug seeking. Thus, repeated use of addictive drugs increases the probability of a cued connection between drug and multiple life events, whereas biologically relevant reinforcers may have fewer cue triggers.

A number of factors are likely to promote drug use over more adaptive behaviors such as the control of sexual behaviors, food and water acquisition, and even complex social behaviors. One is the switch from declarative conscious decision-making behaviors leading to reward to habit-generated behaviors triggered by cues. Another is the repeated increase in dopamine above physiologic levels induced by drug use.

TREATMENT

The treatment of cocaine dependence remains challenging and as yet incomplete compared with the treatment of other addictive disorders. For several addictive disorders, particularly alcohol dependence, a wide array of psychosocial treatments have undergone rigorous scientific study and are known to improve the health of the individual as well as treating the addictive processes. Several US Food and Drug Administration (FDA) medications are available that aid in promoting abstinence, decreasing the likelihood of withdrawal states, and lowering relapse rates. No effective medications have been approved by the US FDA for the treatment of cocaine dependence, although some medicines show promise in early trials, and these are briefly reviewed at the end of this section. There is good empiric evidence for several psychosocial treatments. A recent meta-analysis of 34 well-controlled treatment trials contains much useful information. A total of 9 of the studies were for cocaine dependence, 5 for cannabis dependence, 13 for polysubstance abusers, and the remaining 7 were for opiate dependence. Psychosocial treatments evaluated in these studies included contingency management, relapse prevention, general cognitive behavior therapy, and combinations of cognitive behavior therapy and contingency management. Across this meta-analysis, the average age of the study participants was 34.9 years, 62.2% were male, 61.0% were Caucasian, 67.7% were unmarried, and 42.5% were employed part-time or full-time leaving the majority unemployed. About half of the sample met criteria for

comorbid alcohol dependence. A little over 40% of the study subjects received replacement medication in conjunction with psychosocial treatment. Across all substance use groups, cocaine had the highest mean dropout rate of about 42%. Of the psychosocial treatments, contingency management treatments had the lowest dropout rates at just fewer than 30% followed by general cognitive behavior therapy and combined psychosocial treatments. Efficacy for cocaine-dependent subjects yielded mixed outcomes. Acute abstinence rates were about 32% for cocaine users compared with 13% of subjects in the control treatment conditions. Treatment involving relapse prevention has the largest rate of abstinence at 39%. A higher number of treatment weeks was related to a positive outcome, but the number of treatment sessions per week did not affect treatment outcome or treatment retention. Medication maintenance for opioid agonist therapy did reduce the dropout rate. Limitations of meta-analysis studies include the specific demographics of the sample subjects, the lack of information regarding treatment or treatment as usual, and no information regarding specific comorbid medical and psychiatric conditions.

Self-help support groups abound for the treatment of alcoholism. Large communities have numerous meetings occurring at multiple times every day. Even small rural communities usually have Alcoholics Anonymous meetings on a regular basis throughout the week. Although Narcotics Anonymous and Cocaine Anonymous meetings may be available in large urban areas, they are often not available in medium-sized cities or in small rural communities. Self-help groups have generally not been as intensively studied for the treatment of cocaine dependence as they have been for the treatment of alcohol dependence. Another challenge is that small community-based programs, often run at the level of a county or parish, have limited resources for hiring and training staff, and particularly for using more innovative evidence-based treatment practices, such as contingency management, for their psychosocial treatments.

One of the innovative approaches over the last 20 years has been the development of the Clinical Trials Network (CTN) by the National Institutes of Drug Abuse (NIDA). This is an affiliation of university treatment centers (nodes), which serve as resources to develop affiliation with county, state, and community substance abuse treatment clinics. The CTN nodes serve as a source of funding, innovative training, and clinical trials, some of which can be conducted in community settings. This promotes the innovative transfer of information and skills to smaller treatment centers, establishes scientific methodology for studying program outcomes, and provides clients and CTN-affiliated centers with the latest in effective and/or experimental psychosocial and pharmacological treatments.

Consumer Issues

Clients seeking treatment for cocaine dependence should be encouraged to not only obtain word-of-mouth information regarding the treatment center's reputation, but also avail themselves of resources on the Web about treatment centers using home computers or, when not available, through community centers or public libraries. A treatment center that is part of the CTN, either a university node or a community participation center, is a good sign that the highest level of evidence-based innovative treatments will be used. Another thing to look for is treatment centers that offer multiple services for clients and their families. This is discussed in more detail later. Programs that offer thorough assessments in multiple aspects of the person's medical, social, psychological history are also important. Programs that are individually tailored to a person and their stage of addiction are equally important. An excellent source of information for clients, family members, and professionals is the US Department of Health and Human Services Treatment Improvement Protocol Series published by the Substance Abuse and Mental Health Services Administration Center for Substance Abuse Treatment. Tip 33 (Treatment for Stimulant Use Disorders) is particularly valuable. The Tip Series may be ordered by contacting the National Clearinghouse for Alcohol and Drug Information or through the following Internet site: http://www.kap.samhsa.gov. There is a growing trend for treatment centers to publish their outcome results. Looking at this information, when it is available, is often a helpful guide for clients and their families.

Initial Assessment

Comprehensive treatment of the cocaine-dependent individual begins with a thorough assessment in biological, social, and psychological areas. Many domains of a person's life can be damaged by cocaine, become dysfunctional, or disappear entirely. An initial assessment generally conducted by a multidisciplinary team consisting of physicians, psychologists, social workers, therapists, and nursing personnel are valuable in establishing which parts of a person's life may have been damaged by the ever-growing sphere of influence that cocaine has on a person's life. There is evidence to show that when these services are all available (one-stop shopping), the treatment outcomes are better. However, in some communities, these assessments may need to be gathered piecemeal, and cost may realistically be a barrier to obtaining all of these assessment services in a short period of time. It is generally assumed that a client has the cognitive capacity to participate in any given type of psychosocial treatment, but this may,

at times, be a false assumption. Poor nutrition, concomitant use of other drugs and alcohol, other medical illnesses, head injuries, and educational deficiencies are potential causes of cognitive impairment that may diminish the client's capacity to profit from psychosocial treatments. In most treatment programs, individuals are started in the program, and the therapist detects cognitive problems as they are seen. Then a person may be referred for evaluation of specific cognitive deficits if such skilled personnel are available. If they are not available, switching a client to a form of treatment that can be managed at the patient's current capacity is often necessary. Assessing marital and family problems is also an important part of the process. Vocational problems with unemployment and underemployment also have to be assessed and a practical determination made as to what resources can be brought to bear considering the level of resources that each client may or may not possess. The next section deals with individual forms of treatment. These are somewhat arbitrarily separated for discussion, but in fact, in many treatment centers, these different treatment modalities are intermingled.

Psychosocial Treatment Modalities

The treatment modalities of relapse prevention, community reinforcement, and contingency management have all been studied in double-blind controlled trials and their effectiveness is well established. Some of the other treatment modalities discussed below have not been as well studied, but small research studies and/or clinical consensus support their effectiveness. Many of these therapies have been standardized. Training manuals for therapists and workbooks for clients are available commercially and often through governmental sources such as the Substance Abuse and Mental Health Services Administration and NIDA. This has resulted in greater fidelity of therapies, improved therapist skills, and better client outcomes.

Because return to use/relapse is a ubiquitous feature of all addictions, cognitive behavioral techniques should be applied to develop skills to heighten relapse prevention. Most relapse prevention treatments in use today incorporate elements such as general coping/problem solving, managing slips (light exposure to cocaine) to avoid a major relapse binge, coping with craving, rehearsal of substance refusal and strengthening assertive skills, stress management without drugs and managing seemingly irrelevant decisions and events that frequently led to past relapses. Randomized clinical trials pairing relapse prevention to other psychosocial therapies have not always demonstrated end-of-treatment superiority for relapse prevention techniques. However, the efficacy of relapse prevention strategies appears to strengthen over time. In one study conducted by Carol and colleagues, relapse prevention about as effective as case management procedures at the end of treatment; but at 1 year after treatment, patient outcomes using relapse prevention were superior to case management procedures. Frequently, relapse prevention techniques are now coupled to treatment modalities.

Community reinforcement techniques are often coupled with vouchers so both are discussed together. Community reinforcement is a highly individualized set of treatments designed to promote improvement in several key areas that have been shown to be important for recovery from cocaine dependence. Clients with spouses or significant others are offered marital/couples therapy to repair relationships and build stronger bonds. Unemployed or underemployed clients are given vocational assistance to improve this domain. Usually the group setting of clients' treatment aids in building new social network and recreational behaviors in ways that avoid exposure to cocaine-using friends. To support this latter goal, clients are often encouraged to attend self-help groups such as Cocaine Anonymous and Narcotics Anonymous when such programs are available. Various types of cognitive behavior training build skill proficiency in substance refusal, improve social skills, and provide structure in the person's life and mood control without the use of drugs. When alcohol dependence is a complication of the cocaine dependence, Alcoholics Anonymous self-help groups and medication management such as disulfiram, naltrexone, or acamprosate may be recommended, depending on the health or special needs of individual clients. Voucher-based incentive programs may facilitate abstinence from cocaine and other substances, improve client retention and treatment, and are frequently used as motivational enhancements. Urine drug screen (UDS) monitoring is often conducted two to three times a week and negative UDS results are linked to receiving vouchers. The sophistication of a voucher system can include banking of vouchers with negative urine results and reducing the bank for positive urine results. However, research has noted that a carrot and stick approach with too many aversive consequences leads to higher client dropouts.

Self-Help and 12-Step Groups

Techniques of Alcoholics Anonymous have been scientifically studied in research funded by the National Institutes Alcoholism and Alcohol Abuse. The techniques compare favorably with the cognitive behavioral therapies discussed above, particularly in individual treatment. Large comparison trials enrolling cocaine-dependent clients in therapies such as Narcotics Anonymous or Cocaine Anonymous have not been conducted.

For most clients with cocaine dependence, self-help groups appear to provide an additional source of support with socialization augmenting the formal treatment program. It is important that therapists and other health care providers prepare their clients for self-help groups. When possible, having a client escorted to the first few meetings by an established stable member is a good introduction. Clients need to be reminded that self-help groups may provide contradictory and sometimes even counterproductive information that is in contrast to their standard therapies. This can be particularly problematic for patients with a dual diagnosis and special needs who are on pharmacotherapies for psychiatric disorders. One client with bipolar disorder said, "I was told that being drug free, meant not taking my lithium, too." Re-hospitalization and personal and family suffering can be the result. Therapeutic processing of early self-help group experience is extremely valuable.

Inpatient Treatment

Inpatient treatment programs for cocaine dependence were developed from the 28-day treatment models for alcohol, sometimes called the Minnesota model. Such programs were common in the 1970s and 1980s but have lost favor with private insurance carriers. Treatment in these residential hospital facilities was highly structured and included multiple activities such as 12-step self-help meetings, group psychotherapy, relaxation and coping training classes, educational classes, and individual therapy. Treatment was often supportive and confrontational, combining techniques of motivational training, cognitive behavior therapy, and relapse prevention, but rarely contingency management.

Residential Treatment

Residential treatment programs are usually defined as programs that are 90 days or longer in duration. Such programs are generally reserved for people who fail the intensive outpatient treatments, 28-day treatment programs, or sometimes for individuals whose environment is so destructive to the treatment process that only the distant long-term treatment setting seems plausible. Some residential treatment centers are affiliated with drug courts and receive most of their referrals from the criminal justice system. It is common in these programs for clients to be expected to work at least part-time, and to participate in therapeutic community groups, group psychotherapy, and 12-step self-help meetings. These programs have limited access to medical and psychiatric services on-site. Senior staff members are often clinical psychologists, social workers, or counselors. Much of the therapeutic work structure and guidance are provided by clients who have been enrolled in the program for relatively long periods of time. Residential programs often provide transitions to halfway houses similar to those for patients with other disorders. However, halfway houses for cocaine recovery are often only available in larger communities. Sometimes these are self-regulated group homes where one resident is in charge of rents and ensures the conduct within the community. Clients are expected to work, maintain the rules of the halfway house, stay clean from all drugs, and follow their outpatient treatment.

Treatment as Usual

The term treatment as usual is no more than a cliché that is often used in studies evaluating the efficacy of cognitive behavior therapy, that is, motivational effectiveness training, contingency management, or relapse prevention. Treatment-as-usual programs represent a spectrum of intensity of care and therapist experience and education levels. In larger urban areas, many state programs have the benefit of local state and federal funding plus some fees paid by clients. Intensive outpatient programs are likely to be the norm, sometimes associated with small inpatient detox units that are usually reserved for the treatment of withdrawal from alcohol or sedative hypnotics. Medical and psychiatric services are extremely limited. The exception is methadone maintenance programs, where nursing staff and some physician coverage is provided. In dual-diagnosis programs, psychiatrists and psychologists play significant roles. Rural treatment programs that are underfunded may offer limited individual outpatient counseling services and urine drug screen testing, but empirically-based treatment approaches as described above are often lacking. Twelve-step groups for cocaine-dependent patients may be nonexistent in these areas and travel to self-help groups, group therapy, and other services may represent a significant barrier for clients.

In summary, psychosocial treatments are effective but have significant limitations when viewed from a public health perspective. Even with the best of therapies, only somewhere between a third and a half of individuals attain abstinence from cocaine dependence with any one treatment. Psychosocial treatments are time intensive, expensive, and not available in many parts of the country.

Medication Treatment of Cocaine Dependence

The topic of medication development for the treatment of cocaine dependence is described in detail elsewhere in this volume. In this section, we briefly summarize three biological and pharmacologic areas

that show promise as proof-of-concept therapies augmenting psychosocial treatments for cocaine dependence.

Disulfiram has been used to treat alcohol dependence since the late 1940s. It blocks dopamine beta-hydroxylase in synapses and increases extracellular dopamine levels. Several small double-blind controlled trials in cocaine-dependent subjects have demonstrated efficacy superior to placebo. Nevertheless, questions regarding the safety of disulfiram remain and larger trials evaluating effectiveness and safety have not been conducted. Similar drugs that have more focused pharmacologic activity may be effective and safer.

N-Acetylcysteine is a commonly occurring amino acid used to treat acetaminophen overdose and as a mucolytic agent to treat pulmonary complications of cystic fibrosis and chronic obstructive pulmonary disease. Recent preclinical studies indicate that it is a prodrug converted into cysteine in glial and nerve cells. Extracellular cysteine exchanges one to one with intracellular glutamate to raise extracellular glutaminergic tone and can block cocaine-induced reinstatement and cue-induced reinstatement in animal models. Two small clinical trials, one in nicotine dependence and one in cocaine dependence, indicate promise for this type of therapy. *N*-Acetylcysteine itself is not ideal. It has a relatively short half-life requiring multiple oral administrations during the day, limiting practical compliance. Absorption and bioavailability are poor and more favorable agents are needed to alter glutaminergic tone.

Vaccines that link cocaine to molecules that, when injected into humans, form antigen–antibody complexes that cannot cross the blood–brain barrier is another area in which proof of concept has been established. In human laboratory studies, antibody production is adequate to markedly reduce the subjective experience of smoked cocaine. Not all subjects make adequate antibody levels, the vaccine must be injected frequently, and, with high amounts of cocaine use, the effects of the vaccine can be overridden. It is likely that these technical problems will be overcome.

SUMMARY

Cocaine addiction is a complex phenomenon involving individual vulnerabilities as well as cultural, economic, legal, and demographic modifiers. The clinical course of cocaine dependence varies greatly depending on individuals' vulnerabilities, socioeconomic factors, co-occurrence of psychiatric and medical comorbidities, as well as issues of polysubstance use. In addition, there are social, psychological, and biological factors operating simultaneously to affect the initiation, perpetuation, and course of cocaine dependence.

The treatment of cocaine dependence remains a challenge and is as yet incomplete compared with the treatment of other addictive disorders. Treatments relevant to both clinicians and the lay audience include pharmacologic and psychosocial treatment modalities, initial assessment and inpatient/residential treatment of the individual, and consumer issues including self-help and 12-step groups. Although pharmacologic treatments are limited, disulfiram, *N*-acetylcysteine, and cocaine-binding vaccines are showing promise as potential medications. Psychosocial treatments are currently the most effective but have significant limitations when viewed from a public health perspective. Even with the best of therapies, only somewhere between a third and a half of individuals attain abstinence from cocaine dependence with any one treatment. Psychosocial treatments are time intensive, expensive, and not available in many parts of the country.

SEE ALSO

Gateway Hypothesis, Peer Influences on Addiction

List of Abbreviations

FDA Food and Drug Administration
NIDA National Institutes of Drug Abuse
UDS Urine drug screen

Glossary

Benzoylecgonine the major metabolite of cocaine. Benzoylecgonine is the most common metabolite measured in urine drug screens to detect cocaine use.

Erythroxin coca one of several species of cocaine plants grown in the Andes mountains. Cocaine is contained in the leaves.

***N*-Acetylcysteine (NAC)** an amino acid that is used in clinical medicine to treat acetaminophen overdose and several pulmonary diseases such as cystic fibrosis and emphysema. Some preliminary research indicates that NAC may be of value in treating cocaine and nicotine addictions.

Nucleus accumbens a bilateral collection of neurons forming nuclei that are found within the midbrain. The accumbens plays a major role in reward, learning, addiction, and other compelling behaviors. Although small, the accumbens nuclei have complex functions, morphology, and multiple inputs and outputs.

Further Reading

Brunton, L., Lazo, J., Parker, K. Chapter 23. Drug addiction and drug abuse. In: Brunton, L., Lazo, J., Parker, K. (Eds), The Pharmacological Basis of Therapeutics, Goodman & Gilman.

Calatayud, J., Gonzalez, A., 2003. History of the development and evolution of local anesthesia since the coca leaf. Anesthesiology 98, 1503–1508.

Cox, S.M., Benkelfat, C., Dagher, A., Delaney, J.S., Durand, F., McKenzie, S.A., et al., 2009. Striatal dopamine responses to

intranasal cocaine self-administration in humans. Biological Psychiatry 65, 846–850.
Cramer, S.C., Sur, M., Dobkin, B.H., O'Brien, C., Sanger, T.D., Trojanowski, J.Q., et al., 2011. Harnessing neuroplasticity for clinical applications. Brain 134, 1591–1609.
Dreyer, J.L., 2010. New insights into the roles of microRNAs in drug addiction and neuroplasticity. Genome Medicine 2, 92.
Dutra, L., Stathopoulou, G., Basden, S.L., Leyro, T.M., Powers, M.B., Otto, M.W., 2008. A meta-analytic review of psychosocial interventions for substance use disorders. American Journal of Psychiatry 165, 179–187.
Fields, H.L., Hjelmstad, G.O., Margolis, E.B., Nicola, S.M., 2007. Ventral tegmental area neurons in learned appetitive behavior and positive reinforcement. Annual Review of Neuroscience 30, 289–316.
Ghitza, U.E., Preston, K.L., Epstein, D.H., Kuwabara, H., Endres, C.J., Bencherif, B., et al., 2008. Brain mu-opioid receptor binding predicts treatment outcome in cocaine-abusing outpatients. Biological Psychiatry 68, 697–703.
Gorelick, D.A., Kim, Y.K., Bencherif, B., Boyd, S.J., Nelson, R., Copersino, M.L., et al., 2008. Brain mu-opioid receptor binding: relationship to relapse to cocaine use after monitored abstinence. Psychopharmacology 200, 475–486.
Heard, K., Palmer, R., Zahniser, N.R., 2008. Mechanisms of acute cocaine toxicity. Open Pharmacology Journal 2, 70–78.
Holmes, N.M., Marchand, A.R., Coutureau, E., 2010. Pavlovian to instrumental transfer: a neurobehavioural perspective. Neuroscience and Biobehavioral Reviews 34, 1277–1295.
Huang, Y.H., Schluter, O.M., Dong, Y., 2011. Cocaine-induced homeostatic regulation and dysregulation of nucleus accumbens neurons. Behavioural Brain Research 216, 9–18.
Kalivas, P.W., 2008. Addiction as a pathology in prefrontal cortical regulation of corticostriatal habit circuitry. Neurotoxicity Research 14, 185–189.
Kalivas, P.W., 2009. The glutamate homeostasis hypothesis of addiction. Nature Reviews 10, 561–572.
Kalivas, P.W., O'Brien, C., 2008. Drug addiction as a pathology of staged neuroplasticity. Neuropsychopharmacology 33, 166–180.
Karila, L., Gorelick, D., Weinstein, A., Noble, F., Benyamina, A., Coscas, S., et al., 2008. New treatments for cocaine dependence: a focused review. International Journal of Neuropsychopharmacology/Official Scientific Journal of the Collegium Internationale Neuropsychopharmacologicum (CINP) 11, 425–438.
Karila, L., Reynaud, M., Aubin, H.J., Rolland, B., Guardia, D., Cottencin, O., et al., 2011. Pharmacological treatments for cocaine dependence: is there something new? Current Pharmaceutical Design.
Lodge, D.J., 2011. The medial prefrontal and orbitofrontal cortices differentially regulate dopamine system function. Neuropsychopharmacology 36, 1227–1236.
Lovinger, D.M., 2010. Neurotransmitter roles in synaptic modulation, plasticity and learning in the dorsal striatum. Neuropharmacology 58, 951–961.
Lynd-Balta, E., Haber, S.N., 1994. The organization of midbrain projections to the ventral striatum in the primate. Neuroscience 59, 609–623.
Nestler, E.J., 2005. The neurobiology of cocaine addiction. Science & Practice Perspectives/a publication of the National Institute on Drug Abuse, National Institutes of Health 3, 4–10.
Olds, J., Milner, P., 1954. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. Journal of Comparative and Physiological Psychology 47, 419–427.
Olsen, C.M., 2011. Natural rewards, neuroplasticity, and non-drug addictions. Neuropharmacology.
Rastegar, D., Fingerhood, M., 2005. Alcohol. Addiction Medicine: An Evidenced-Based Handbook. Lippincott Williams & Wilkins, Philadelphia, PA, pp. 36–75.
Sellman, D., 2009. The 10 most important things known about addiction. Addiction (Abingdon, England) 105, 6–13.
Sesack, S.R., Grace, A.A., 2010. Cortico-basal ganglia reward network: microcircuitry. Neuropsychopharmacology 35, 27–47.
Sofuoglu, M., Kosten, T.R., 2006. Emerging pharmacological strategies in the fight against cocaine addiction. Expert Opinion on Emerging Drugs 11, 91–98.
Sofuoglu, M., Poling, J., Gonzalez, G., Gonsai, K., Kosten, T., 2006. Cocaine withdrawal symptoms predict medication response in cocaine users. American Journal of Drug and Alcohol Abuse 32, 617–627.
Stuber, G.D., Hopf, F.W., Tye, K.M., Chen, B.T., Bonci, A., 2010. Neuroplastic alterations in the limbic system following cocaine or alcohol exposure. Current Topics in Behavioral Neurosciences 3, 3–27.
Thomas, M.J., Kalivas, P.W., Shaham, Y., 2008. Neuroplasticity in the mesolimbic dopamine system and cocaine addiction. British Journal of Pharmacology 154, 327–342.
US Department of Health and Human Services, 2006. Treatment for Stimulant Use Disorders. Treatment Improvement Protocol (TIP) Series 33. Substance Abuse and Mental Health Services Administration.
Volkow, N.D., Wang, G.J., Fischman, M.W., Foltin, R.W., Fowler, J.S., Abumrad, N.N., et al., 1997. Relationship between subjective effects of cocaine and dopamine transporter occupancy. Nature 386, 827–830.
Volkow, N.D., Wang, G.J., Telang, F., Fowler, J.S., Logan, J., Childress, A.R., et al., 2006. Cocaine cues and dopamine in dorsal striatum: mechanism of craving in cocaine addiction. Journal of Neuroscience 26, 6583–6588.
Wong, D.F., Kuwabara, H., Schretlen, D.J., Bonson, K.R., Zhou, Y., Nandi, A., et al., 2006. Increased occupancy of dopamine receptors in human striatum during cue-elicited cocaine craving. Neuropsychopharmacology 31, 2716–2727.

Relevant Websites

http://www.nida.nih.gov/nidahome.html–National Institute of Drug Abuse;
http://www.samhsa.gov/–The Substance Abuse and Mental Health Service Administration.

CHAPTER

69

Marijuana Use and Abuse

*Kristin W. Grover**, *Michael J. Zvolensky**, *Marcel O. Bonn-Miller*[$]*, *Jesse Kosiba**, *Julianna Hogan**

[*]The University of Vermont, Burlington, VT, USA, [$]National Center for PTSD and Center for Health Care Evaluation, VA Palo Alto Health Care System, Palo Alto, CA, USA

OUTLINE

INTRODUCTION

Marijuana (also referred to as cannabis) is a drug derived from the flowers, stems, leaves, and seeds of the hemp plant (*Cannabis sativa*). The need for public health awareness and evidence-based clinical care for marijuana use and its disorders remains a major health care priority in the United States and beyond. Rates of marijuana use, abuse, and dependence in the United States represent a significant public health concern considering that several well-documented negative consequences have been associated with daily or weekly drug use (e.g. increased risk of severe medical disease, increased risk-taking behavior, and clinically significant life impairment).

The main aim of the present chapter is to provide an overview of marijuana use and its disorders. The chapter is organized into seven sections. First, we describe the prevalence of marijuana use and its disorders. Second, we clarify the nature of marijuana use in terms of its pharmacokinetics and acute intoxication features. In the third section, we describe the classification of marijuana use and its disorders using the current diagnostic nomenclature. Fourth, we describe the motivational bases for use of the drug. In the fifth section, we provide a synopsis of some problems associated with marijuana use and its disorders, including health problems, social problems, and psychological disturbances. Sixth, we provide a summary of the scientific work focused on marijuana, the reasons for its use, and users' relative success in quitting. In the final section, we describe some practically oriented clinical issues for primary care medical practitioners to consider in terms of the recognition and treatment of marijuana use and its disorders.

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00069-3

PREVALENCE

Marijuana has been the most widely used illicit substance in the United States for 30 consecutive years, with approximately 25 million Americans (8.6% of the population) having used marijuana in the past year, and approximately 15.2 million (6.1% of the population) having used marijuana in the past month. An estimated 10% of individuals who have ever used marijuana will become daily users. Lifetime marijuana dependence is estimated at 4% of the general population, a rate that is the highest of any illicit drug. Rates of conditional dependence, defined as the risk for developing dependence among those who have ever used the drug, indicate that marijuana is associated with a high rate of dependence potential. For example, the relative risk of marijuana dependence, given use of the drug in the past year, is estimated to be 7% among adults, which is only slightly lower than that for cocaine (12%) and greater than that observed for alcohol (5%). Furthermore, greater levels of use are related to an increased risk for dependence. Studies suggest that the rate of dependence is 20–30% among persons using marijuana on a regular (weekly) basis. Moreover, marijuana use problems have increased, with 35% of adult marijuana users in the United States currently meeting criteria for marijuana abuse or dependence, compared with 30% 10 years earlier, representing an increase of approximately 730 000 individuals.

Treatment and community studies have examined prevalence rates of marijuana use among samples suffering from a variety of medical and psychological problems. Among those seeking treatment for psychosis, approximately 23% currently use marijuana, with about half of that group currently misusing the drug. It has been reported that approximately 16% of people with spinal cord injury use marijuana. Other work has found that marijuana use accounts for as much as 25% of the primary drug problems of individuals seeking residential drug treatment. These studies suggest that marijuana use may be overrepresented among certain vulnerable populations.

PHARMACOKINETICS AND INTOXICATION FEATURES

Pharmacokinetics

Marijuana can be consumed via smoking (e.g. hand-rolled cigarettes, water pipes, nonwater pipes) or ingestion (e.g. mixed into foods, used in the process of brewing tea). The active agents in marijuana are cannabinoids (unique to the marijuana plant). There are at least 60 different cannabinoids in marijuana, although the pharmacokinetics of the vast majority of these compounds is largely unknown. Of these, the most well-known, and arguably the most important cannabinoid is tetrahydrocannabinol, which is believed to be the most potent psychoactive agent in the cannabinoid plant. The tetrahydrocannabinol content of plants from a range of sources and strains varies dramatically. With a focus on improved plant breeding and improved growing techniques, the tetrahydrocannabinol content of marijuana has increased dramatically. As one illustrative example, tetrahydrocannabinol content from a typical marijuana cigarette (joint) in the 1960s was 10 mg, whereas estimates suggest that it is currently around 1 g (or 150–200 mg). Given that the effects of marijuana are dose dependent (i.e. greater amount or potency yields greater effect), the significantly increased potency of marijuana is a major public health concern and is important in understanding the current and historical prevalence rates of use, abuse, and dependence.

Since the discovery of a cannabinoid receptor within the brain in the late 1980s, researchers have been able to explicate the process by which tetrahydrocannabinol acts on the brain. Currently, there is evidence of three potential cannabinoid receptors, only one of which is located within the brain (the cannabinoid-1 receptor). When tetrahydrocannabinol is inhaled into the body via marijuana smoking, it passes from the lungs into the bloodstream. Once in the blood, tetrahydrocannabinol attaches to cannabinoid receptors, such as the cannabinoid-1 receptor, adding to or reducing the naturally occurring endogenous ligands for these receptors (e.g. anandamide). The cannabinoid-1 receptor, in particular, has been found to mediate both neurochemical and behavioral properties of these cannabinoids including tolerance. It also is noteworthy that tetrahydrocannabinol and other cannabinoids move rapidly into fat and other body tissues but are slowly released from these tissues back into the bloodstream. Eventually, cannabinoids are cleared from the body via urine and fecal matter.

Acute Intoxication Features

In general, marijuana consumption produces a mild, relatively short period of intoxication (being high). Specifically, marijuana can produce a range of acute psychosensory experiences including perceptual distortions (e.g. hallucinogenic properties), relaxation, anxiety, acute paranoia, inhibition, and so on. Periods of intoxication depend on use patterns and potency, but tend to last for at least a few hours. Marijuana intoxication also impairs cognitive and psychomotor performance with complex, demanding tasks. There is a dose-dependent relation between marijuana use and psychomotor and

cognitive impairment, with higher doses being associated with more impairment for more demanding tasks. Although cognitive impairment for hours after using marijuana is a well-replicated phenomenon in laboratory studies, there has been consistent debate about the permanent cognitive effects of using marijuana. Some recent work suggests that individuals who have used marijuana over long periods of time demonstrate impaired performance on a variety of neuropsychological tests (e.g. attention, memory, and processing complex information) even when not acutely intoxicated. These negative cognitive effects appear to be present months and even years after successful cessation.

CLASSIFICATION

The current diagnostic criteria for problematic patterns of marijuana use, according to the *Diagnostic and statistical manual of mental disorders, 4th edition* (DSM-IV-TR), include abuse and dependence (see Table 69.1 for the diagnostic criteria for marijuana abuse and dependence). Marijuana abuse is a pattern of marijuana use that includes significant and unpleasant consequences associated with frequent use within a 12-month period. Some of the consequences associated with marijuana abuse include multiple legal problems, repeated use in physically hazardous situations, and recurrent social and interpersonal problems as a result of use. What differentiates substance abuse from dependence is that abuse only includes harmful consequences of frequent use, whereas dependence indicates compulsive use, tolerance, or withdrawal. As with diagnosis of other substance use disorders, marijuana abuse cannot be diagnosed if marijuana dependence criteria are met, highlighting the putative, more severe nature of marijuana dependence.

There are relatively few empirical data pertaining to the validity of distinguishing among marijuana use, abuse, and dependence. Moreover, for a long period of time, scholars did not uniformly endorse or support a marijuana dependence syndrome. Current research has partially laid these earlier questions to rest in that heavy users of the drug tend to report problems controlling their use, despite noted negative consequences, and experience withdrawal and other adverse symptoms when discontinuing use. In fact, estimates suggest that approximately one out of every ten individuals who use the drug become dependent on it at some point in time.

To date, researchers have used standardized interviews to index marijuana diagnoses in a manner identical to those for other types of substances (e.g. alcohol, tobacco). In contrast to the relatively recent emerging perspective that classification of marijuana along the nosological lines of use, abuse, and dependence is the optimal and most accurate approach, it has been more common historically to denote marijuana use variability by asking respondents to indicate their level of use (e.g. frequency) over a specified period of time. Collectively, then, deciding on whether nosological classification and/or a use-oriented assessment protocol (i.e. volume and frequency) is indicated may depend on the specific clinical need or research question being posed and the theoretical basis for it.

MOTIVATIONAL BASES

Researchers and clinicians have increasingly found merit in applying motivational models to understand and clinically intervene with marijuana use and its disorders. At the most basic level, such an approach recognizes that there are a number of distinct motives for using marijuana that can vary both between and within individuals. Motivational models predict that distinct motives may theoretically be related to particular types of problems. For example, specific motives may play unique roles in various aspects of use (e.g. addictive use, withdrawal symptoms, craving) or problems related to use (e.g. psychological disturbances, risk-taking behavior). Thus, enhancing efforts to explicate marijuana use motives empirically will presumably facilitate our understanding of the nature of marijuana use and its disorders as well as linkages between marijuana use and its clinically important correlates.

The Marijuana Motives Measure was developed to assess marijuana use motives. This measure has demonstrated a multidimensional measurement model across extant work – specifically, a five-factor solution denoting Enhancement, Conformity, Expansion, Coping, and Social motives for marijuana use, each with satisfactory levels of internal consistency. Existing motivation-oriented work on marijuana is important in terms of informing the understanding of how and why marijuana use may be related to certain patterns of substance use and psychological problems. Greater levels of Coping, Enhancement, Social, and Expansion motives for marijuana use have each been found to be concurrently, significantly associated with frequency of marijuana use in the past 30 days. Likewise, more severe forms of marijuana use (e.g. dependence) are associated with greater motivation to use marijuana for multiple reasons. These associations between motives for use and frequency of use do not appear to be attributable to alternative factors such as the amount of time being a marijuana user or other types of concurrent substance use. However, the exact directional relation between marijuana motives and patterns of marijuana use remains underexplored. Still, it is noteworthy that specific motives may be relevant to the understanding of psychological vulnerability.

TABLE 69.1 Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision Criteria for Marijuana Abuse and Dependence

Abuse criteria

A. Maladaptive pattern of substance use, leading to clinically significant impairment or distress, as manifested by one (or more) of the following, occurring within a 12-month period:

- **I.** Recurrent use resulting in failure to fulfill major role obligations at work, school, or home
- **II.** Recurrent use in which it is physically hazardous
- **III.** Recurrent marijuana-related legal problems
- **IV.** Continued use despite persistent or recurrent social or interpersonal problems caused by or exacerbated by the effects of marijuana

B. The symptoms have never met criteria for marijuana dependence

Dependence criteria

A. Maladaptive pattern of substance use, leading to clinically significant impairment or distress, as manifested by three (or more) of the following, occurring at any time in the same 12-month period:

- **I.** Tolerance, as defined by either of the following:
  - **(a)** A need for markedly increased amounts of the substance to achieve intoxication or desired effect
  - **(b)** Markedly diminished effect with continued use of the same amount of the substance
- **II.** Withdrawal, as defined by either of the following:
  - **(a)** Characteristic withdrawal syndrome associated with marijuana
  - **(b)** The same (or closely related) substance is taken to relieve or avoid withdrawal symptoms (Note: In the current version of DSM-IV-TR, marijuana withdrawal is not a diagnostic criterion)
- **III.** Marijuana is taken in greater amounts or over a longer period of time than was intended
- **IV.** There is a persistent desire or effort to cut down or control marijuana use
- **V.** A great deal of time is spent in activities necessary to obtain the substance, use the substance, or recover from its effects
- **VI.** Important social, occupational, or recreational activities are given up or reduced because of marijuana use
- **VII.** The substance use is continued despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by marijuana

For example, coping motives for marijuana use, but not other motives, have been significantly predictive of negative affect, anxious arousal, posttraumatic stress symptoms, and anhedonic depressive symptoms. These types of findings may have important theoretical implications for a better understanding of previous research linking marijuana use to affect-based psychological vulnerability.

NEGATIVE CORRELATES

Marijuana has historically been viewed by some as a less severe or soft drug. In contrast, scientific study has provided empirical evidence that marijuana use and its disorders are associated with a number of clinically significant problems. Indeed, there are several empirically documented negative consequences of frequent or problematic marijuana use (typically defined as weekly or daily use). These negative effects are evident in physical, social, interpersonal, and, more recently, psychological realms. In this section, we describe examples of work pertaining to possible negative correlates of marijuana use.

Health-Related Problems

Perhaps the foremost negative effect linked to various types of marijuana use is its impact on physiological processes, particularly the cardiovascular and pulmonary systems. For example, frequent marijuana use is associated with increased risk of severe respiratory illnesses, especially chronic bronchitis. Other work has shown that compared with individuals who do not use marijuana or tobacco, or with tobacco smokers who have no history of marijuana use, the lung function of those who use marijuana regularly is significantly poorer. A series of important large-scale prospective studies also has documented the negative effects of marijuana over time on pulmonary functioning. Though

the results across investigations are not fully consistent, they converge on the observation that greater duration of marijuana use is related to increased bronchitis symptoms (e.g. coughing, wheezing).

There also is a link between marijuana use and cancer. Most investigations suggest that there is an increased risk of lung cancer among more frequent and longer-term users of the drug. Controlled studies of these cancer-related negative effects of marijuana use, however, are largely underrepresented in the literature. In addition to the increased risk for lung cancer, it is noteworthy that some research suggests that marijuana use may be related to impaired immune system functioning, but these investigations, again, have not been consistently replicated. Upon close inspection of these studies, it becomes clear that some of the inconsistencies may be related to problems in the measurement of marijuana use and individual differences in use. A similar set of issues is evident for links between marijuana use and impaired reproductive effects. Nonhuman research suggests that heavier marijuana use is related to impaired reproduction capacity, but controlled evidence in humans is currently lacking.

It should be noted that although the vast majority of research has focused on elucidating putative negative health consequences or correlates of marijuana use, there has been scientific and clinical interest on possible health benefits of the drug. Namely, it has been suggested that marijuana improves certain disease symptoms (e.g. decrease in perceived pain, decrease in eye pressure, decrease in involuntary movement) and stimulates appetite. Although this body of work is complicated, the strongest evidence of possible health benefits for marijuana use appears to be focused on increasing appetite, decreasing nausea and vomiting, preventing systemic weight loss, and possibly improving pain tolerance. Notably, a recent meta-analysis including 18 controlled trials that compared cannabis treatment to placebo concluded that cannabis treatment is moderately effective for treating chronic pain. However, this analysis also indicated that there are a number of potentially harmful effects associated with cannabis treatment (e.g. impaired cognitive and motor function, mood disturbances). These findings therefore suggest that this treatment may be potentially more risky than beneficial for patients with chronic pain in the long term.

Social Problems

In addition to a number of negative physical consequences, adverse social consequences related to frequent marijuana use have been reported. For example, marijuana use is a contributing factor to impaired educational attainment, reduced workplace productivity, and impaired judgment within hours after marijuana use (e.g. among airline pilots). In all of these studies, a consistent pattern emerges: the greater the amount of use (measured in frequency of use or severity of use), the greater the impairment. The specific mechanism(s) underlying these use-related effects are as yet theoretically and empirically unspecified.

Marijuana use also has been shown to be related to other social problems. For example, quantity of marijuana use and acute intoxication have been related to general risk-taking behavior and impaired judgment. For instance, marijuana use has been linked to fatal traffic accidents and general driving impairment, even after statistically controlling for the variance accounted for by alcohol use. Also evident of risk-taking behavior, marijuana use has been found to be associated with risky sexual behavior among young adult women. Other work suggests that frequent or more severe marijuana use may lead to using more severe forms of other drugs (e.g. gateway theories). One overarching limitation to work linking certain types of marijuana use to social and interpersonal functioning, and even future use of other substances, is that there is a dearth of prospective evaluations.

Psychological Problems

A variety of psychological problems seem to be associated with marijuana use and its disorders, with psychotic spectrum disorders being more prominent. Indeed, case reports have documented that marijuana use can precede the onset of psychotic spectrum disorders such as schizophrenia at higher rates than what would be expected by chance of psychosis among regular marijuana users. Although the directional nature of the marijuana and psychotic spectrum problem association has been the subject of debate, one position has been that the use of marijuana may actually increase the risk of psychotic spectrum disorders. Consistent with this marijuana to psychotic symptoms/disorders perspective, the acute effects of marijuana use have been found to contribute to psychotic episodes and exacerbations of such symptoms among previously afflicted persons (e.g. the recurrence of psychotic symptoms). Other work has found that intravenous tetrahydrocannabinol administered to patients with schizophrenia treated with antipsychotics and nonpsychiatric controls exacerbated positive schizophrenic symptoms in the patient sample and induced positive symptoms in controls. Neuroimaging studies have found similarities between neural networks impaired by marijuana use and those known to be implicated in the cause of schizophrenia. Finally, two meta-analytic reviews of the existing empirical literature both concluded that the use of marijuana increased the risk of schizophrenia and psychotic symptoms and disorders. Although a model

proposing that marijuana may lead to psychotic spectrum disorders provides only one possible way in which these factors may be related, it documents the importance of understanding marijuana in the context of severe mental illness.

Another area of research has focused on addressing the relationship between marijuana and depressive symptoms or problems. Clinically, regular (e.g. on a daily or weekly basis) marijuana users often report a lack of motivation for completing day-to-day activities (e.g. going to school). The literature on depression and marijuana has sometimes identified statistically significant relations between marijuana use and depressive symptoms and disorders. However, the strength of the associations between marijuana and depression may be relatively weak after controlling for common variables such as gender. As one illustrative example, it has been found that marijuana use in early adolescence does not significantly predict later depressive symptoms after controlling for distress and interpersonal functioning in earlier adolescence. These findings, when considered in the context of the psychotic spectrum research noted earlier, highlight that marijuana should not be considered to have the same types of linkages with all forms of mental illness.

Another stream of more recent work has begun to address the relations between marijuana use and anxiety symptoms and disorders. This work was initially stimulated by the observation that marijuana use may acutely promote heightened levels of anxiety symptoms and elicit panic attacks under certain conditions or among certain individuals. For example, approximately 40% of weekly marijuana users reported at least one panic attack related to such use. These prevalence rates are noteworthy in light of lifetime rates of panic attacks among the general population of approximately 5–8%. The interaction between marijuana use and anxiety sensitivity (fear of anxiety and related internal sensations) is related to increased levels of anxiety symptoms among marijuana users who also use tobacco. Thus, certain individual differences such as anxiety sensitivity may be important to consider in understanding the links between marijuana use and anxiety states and disorders.

It also has been found that a lifetime history of marijuana dependence, but not use or abuse, is related to an increased risk of panic attacks after covarying the effects of polysubstance use, alcohol abuse, and demographic variables. Likewise, other research also has found that adolescent-onset marijuana use and dependence are significantly, prospectively associated with increased odds for the development of panic attacks and panic disorder. However, marijuana use or dependence is not incrementally associated with the development of panic after controlling for daily cigarette smoking. These recent findings underscore the importance of considering the role of cigarette smoking in the context of marijuana use in regard to understanding panic vulnerability.

Another emerging area of study has focused on marijuana and posttraumatic stress disorder (PTSD). Extant, albeit limited, empirical work has documented that (1) PTSD symptoms are associated with greater frequency of marijuana use; (2) greater (subclinical) PTSD symptom severity among persons exposed to trauma is significantly related to increased coping-oriented marijuana use (e.g. "To forget my worries"); and (3) hyperarousal symptoms, generally, and sleep problems, specifically, among those with PTSD may be key mechanisms underlying coping-oriented marijuana use.

Recent research indicates that coping motives may facilitate the relation between various anxiety constructs and marijuana use and disorders. Specifically, coping motives for marijuana use mediate the relation between anxious arousal symptoms and frequency of current marijuana use, as well as the relation between anxiety sensitivity and marijuana dependence. This research indicates that people may use marijuana as a means of coping with anxious arousal symptoms and anxiety sensitivity.

MOTIVATION TO QUIT, REASONS FOR QUITTING, AND SUCCESS IN QUITTING

It is important to point out that marijuana has many cardinal features of addiction similar to hard drugs. For many individuals who use marijuana, tolerance to the drug develops and, presumably, contributes to more frequent or heavier use patterns or dosing with more potent (more pure tetrahydrocannabinol) forms of the drug. For example, nonhuman research and, more recently, a smaller human empirical database suggest that marijuana discontinuation among regular users produces an internally consistent withdrawal pattern. Disrupted sleep, nightmares, nausea, anxiety, tension, irritability, sweating, and chills are common withdrawal symptoms. Many of these symptoms appear early after drug discontinuation, and some may last for weeks beyond the quit day (e.g. disrupted sleep). This withdrawal profile can appear relatively quickly during the course of use (e.g. relatively early in the marijuana using career) and may have clinical importance in terms of predicting relapse, although current data are not yet sufficiently developed to yield conclusions in this regard. With the recognition that marijuana use and its disorders are common addictive behaviors and can be related to life impairment and a variety of related negative consequences, it is natural to question how motivated users are to quit, what their reasons are for

quitting, and what their relative degree of success is in doing so.

Motivation to Quit

Two bodies of empirical evidence indicate that a large number of individuals who use marijuana on a regular basis (e.g. monthly) and who meet a range of diagnostic criteria (from use to dependence) are motivated to quit. The first literature has evaluated treatment-seeking behavior. Here, the Drug Abuse Reporting Program and other reports first documented that a clinically significant number of individuals were seeking therapeutic services for problematic marijuana use. Other large-scale surveys independently replicated such findings. For example, 35% of adults admitted to the US public treatment system in 1998 were admitted for treatment of marijuana problems. This rate is higher than those found for cocaine (32%), opioids (18%), stimulants (9%), and other psychoactive substances (12%). In addition, other reports involving national databases have found that the demand for treatment of marijuana use and its disorders doubled between 1992 and 1998. Even college students, who are often perceived as lacking interest in treatment, report interest in marijuana treatment. For example, studies have found that nearly one-fourth of students with more than one marijuana-related problem expressed interest in marijuana treatment, and nearly 60% of adolescent marijuana users indicated that they would participate in a marijuana treatment program targeting adolescents. It also is important to note that marijuana treatment outcome studies have documented that a large number of treatment-seeking marijuana users are not current polysubstance abusers. For example, 80% of a large, marijuana-dependent sample did not report abuse of other substances in the past 90 days and 40% reported never abusing an illicit drug other than marijuana. These data indicate that marijuana represents a significant clinical and public health problem in its own right and commonly prompts treatment-seeking behavior even in the absence of other drug use.

The second body of evidence related to motivation to quit suggests that, despite the notable rates of documented treatment-seeking behavior, most persons using, abusing, or dependent on marijuana actually attempt to quit on their own. Self-quit behavior is operationally defined as attempts to quit without professional assistance (i.e. enrolling in a formal treatment program that uses pharmacological, psychosocial, or combined therapeutic approaches). Numerous studies have reported that by young adulthood, many individuals have made multiple marijuana quit attempts on their own. It also is noteworthy that rates of self-quit attempts from marijuana are generally similar to those observed for other substances (e.g. tobacco). For example, studies of weekly marijuana users have indicated that by age 30, individuals have reported a range of three to seven quit attempts on their own. Although some of these unsuccessful quitters may seek professional treatment when they continue to fail in their quit efforts, it is not presently clear what percentage will ultimately do so and under what circumstances.

Reasons for Quitting

Current marijuana users, ranging from monthly users to those dependent on the drug, report multiple concurrent reasons for quitting. Among adults, worry about physical and psychological effects of marijuana use is the factor for wanting to quit most often cited. For example, 60% of non–treatment-seeking adult weekly marijuana users reported worry about health problems (both real and perceived) as a motivating factor for quitting, and 63% desired to quit in order to gain more self-control over their lives. Similarly, other research has found that anxiety or depressive symptoms are the most commonly reported negative effects of marijuana use and the primary reason for quitting among weekly marijuana users. Others have reported similar findings among both nontreatment seekers and treatment seekers. Such findings do not seem to vary as a function of the type of marijuana use problem. Overall, these data suggest that marijuana users typically express multiple reasons for quitting, with the most common reasons pertaining to excessive negative emotional symptoms (and impaired levels of personal self-control associated with regular marijuana use).

Success in Quitting

Individuals attempting to quit marijuana experience marked difficulty whether they make a quit attempt on their own or seek professional treatment. Numerous survey studies, for example, have documented that current, regular marijuana users (both those who are and are not dependent on the drug) who try to quit on their own report difficulty in remaining abstinent, as indexed by numerous unsuccessful quit attempts.

Although self-quit attempts (without professional assistance) tend to be the most frequently used quit strategy, it is striking that even among those who do seek professional treatment, relapse to use is a common experience. Indeed, a critical review of the treatment outcome literature for marijuana dependence concluded that patients do not show a positive response to treatment, suggesting that it is difficult to treat marijuana dependence. For example, 63% of adults receiving two

of the best available intervention strategies – motivational individualized intervention or cognitive behavioral therapy – relapsed to regular use within 4 months. At 16 months, relapse rates rose to 71 and 72% for the motivational individualized intervention and cognitive behavioral therapy, respectively. Other studies have reported similar results, and more recent clinical trials have extended such work by noting that in addition to full relapse, lapses are very common and clinically significant. For example, among marijuana-dependent adult outpatients receiving treatment, 71% lapsed (defined as any marijuana use) within 6 months, 46% within 3 months, and 24% within 1 month. In addition, 71% of lapsers ultimately experienced a full relapse (defined as four or more days of use per week).

It also should be noted that there have been historically few pharmacotherapy options available for marijuana use disorders. In fact, currently there are no medications approved by the United States Food and Drug Administration for marijuana use disorders, although a number of agents are currently being investigated.

Although marijuana relapse is now a well-documented, prevalent clinical problem, there has been relatively little scientific work focused on predictors of success or failure in attempts to quit using marijuana. The work that has been completed has been broadly guided by social learning, stress and coping, and behavioral economic theories of substance use and relapse. These studies have thus far provided a number of initial and important observations: (1) early lapses are predictive of later relapses among adult and adolescent individuals abusing or dependent on marijuana, regardless of whether they receive formal treatment or not; (2) personal stressors (e.g. family conflict) are related to relapse among individuals who abuse or are dependent on marijuana and are receiving outpatient treatment; (3) other substance use and substance use by peers are predictive of relapse to marijuana use among adolescent outpatients who are abusing or dependent on marijuana; and (4) the level of self-efficacy (i.e. beliefs regarding one's ability to refrain from use) for abstaining from marijuana use among adults who are abusing or dependent on marijuana and are seeking treatment is predictive of later relapse.

CLINICAL ISSUES

Given that marijuana use and its disorders are common and can be associated with a relatively wide variety of negative problems, clinicians, such as primary care physicians who interact with patients in nonspecialty clinical settings, ought to be knowledgeable about the basic issues in clinical care for this drug problem. To facilitate this process, we now turn to a discussion of some core clinical competencies by highlighting basic assessment and treatment strategies. This discussion is broadly relevant to clinical practitioners working in medical, dental, and psychological sectors of the health care industry.

Basic Competencies

The most basic level of competency of clinical relevance focuses on simply being aware of the scientifically developed knowledge on the prevalence and impact of marijuana and its disorders. Clinicians should initially strive to attain an overall awareness of marijuana use and behavior as it relates to their patient population(s). By obtaining such knowledge, the clinician is better equipped to offer patients accurate information about problems related to marijuana use. More specific goal-oriented targets can include, but are not limited to, being able, efficiently and capably, to: (1) describe the prevalence of marijuana use and its disorders; (2) describe regional marijuana use patterns; (3) describe the negative physical and psychological consequences of marijuana use and dependence; (4) describe the importance and role of marijuana treatment, particularly those methods based on evidence-based resources; (5) maintain a general awareness of emerging research related to the treatment of marijuana use and its disorders; (6) understand the criteria used for defining marijuana use, abuse, and dependence; and (7) communicate an interest and willingness to consult with other resources when marijuana knowledge may be limited.

A second basic competency skill domain pertains to developing basic assessment and counseling skills for dealing effectively with marijuana use and its disorders. This domain of competence builds naturally from the foregoing description of general knowledge and awareness. This area of work necessarily begins with developing a level of clinical comfort with marijuana use topics and being capable of engaging a patient in a discussion focused on this topic. From the counseling perspective, a variety of core skills are necessary, including, but not limited to: (1) having the capacity to be an active listener and demonstrate an empathetic stance regarding clinical care involving marijuana-related issues; (2) being able to communicate the strengths and challenges of evidence-based care treatment approaches for marijuana use and its disorders in a nonthreatening manner; and (3) being able to understand basic models of behavior change that pertain to marijuana use and meaningfully communicate levels of motivational stage and readiness to clients.

From an assessment perspective, basic competencies are needed to understand how to evaluate marijuana

use behavior and history adequately. Without this level of proficiency, it will be challenging to document readiness to quit or success in doing so. The overarching goal is to learn to comprehensively document and obtain accurate information that can be used in a clinically meaningful manner. The assessment process can be usefully divided into two global phases: intake (or initial assessment) and ongoing assessment. For the intake assessment, key variables to assess include the extent and nature of marijuana use from a lifetime and current perspective; documenting current interest and motivation in quitting; using evidence-based technologies for documenting marijuana use, abuse, and dependence; identifying (with the client) barriers to quitting currently; identifying strengths in the client or the environment (e.g. social support) for quitting; documenting the nature of past quit history and relative degree of success in such attempts; and personal as well as cultural variables that may affect marijuana use and decisions regarding use. The intake assessment process should also integrate information about the client's medical and psychological history in order to understand how such factors may influence the ongoing marijuana use or attempts to quit.

Ongoing assessments require an understanding of each client and the specific variables that need to be regularly tracked to accurately and objectively document (and understand) the motivation to quit marijuana use behavior. Marijuana use behavior, ongoing life stressors, and current motivation to quit are possibly important targets. This information can be used to track and understand ongoing efforts to quit, and to help clients formulate a plan for making a quit attempt that is individualized to their specific needs and life circumstances.

Aside from the individual level of commitment to professional development, it is a reality that most medical care occurs within a context that intersects with other health care professionals. Therefore, enlisting the systems involved in such clinical work in an integrated manner may be a powerful resource for dealing with marijuana use and its disorders. The need for such systems-oriented care is particularly evident given that educational efforts solely focused on the individual have not always been met with large degrees of success in the substance use field.

SUMMARY

Understanding and treating marijuana use is an important public health priority. Despite the increasing recognition that marijuana use and its disorders are not actually harmless, the scientific literature pertaining to the cause and maintenance, assessment, and treatment of marijuana use and its disorders is still in its beginnings. The next decade promises to be an important time to marshal resources to bridge major knowledge gaps and translate such developments into promising prevention and treatment approaches.

List of Abbreviations

DSM-IV-TR	diagnostic and statistical manual of mental disorders, 4th edition, text revision.
PTSD	posttraumatic stress disorder.

Further Reading

Bloom, J.W., Kaltenborn, W.T., Paoletti, P., Camilli, A., Lebowitz, M.D., 1987. Respiratory effects of non-tobacco cigarettes. British Medical Journal 295, 1516–1518.

Bonn-Miller, M.O., Vujanovic, A.A., Feldner, M.T., Bernstein, A., Zvolensky, M.J., 2007. Posttraumatic stress symptom severity predicts cannabis use coping motives among traumatic event-exposed cannabis users. Journal of Traumatic Stress 20, 577–586.

Budney, A.J., Higgins, S.T., Radonovich, K.J., Novy, P.L., 2000. Adding voucher-based incentives to coping skills and motivational enhancement improves outcomes during treatment for marijuana dependence. Journal of Consulting and Clinical Psychology 68, 1051–1061.

Budney, A.J., Moore, B.A., Vandrey, R., Hughes, J.R., 2003. The time course and significance of cannabis withdrawal. Journal of Abnormal Psychology 112 (3), 393–402.

Compton, W.M., Grant, B.F., Colliver, J.D., Glantz, M.D., Stinson, F.S., 2004. Prevalence of marijuana use disorders in the United States: 1991–1992 and 2001–2002. Journal of the American Medical Association 291, 2114–2121.

Hall, W., Solowij, N., 1998. Adverse effects of cannabis. Lancet 352, 1611–1616.

Stephens, R.S., Roffman, R.A., Curtin, L., 2000. Comparison of extended versus brief treatments for marijuana use. Journal of Consulting and Clinical Psychology 68, 898–908.

Zimmer, L., Morgan, J.P., 1997. Marijuana Myths, Marijuana Facts. The Lindesmith Center, New York.

Relevant Websites

http://www.nida.nih.gov/researchreports/marijuana/marijuana3.html – National Institute on Drug Abuse: Marijuana Abuse.

http://ncadi.samhsa.gov/ – Substance Abuse and Mental health Services Administration.

http://www.safeaccessnow.org/ – Americans for Safe Access: Advancing Legal Medical Marijuana Therapeutics and Research

http://www.maps.org/ – Multidisciplinary Association for Psychedelic Studies

CHAPTER

70

Methamphetamine Addiction

*Rebecca McKetin**, *Sharlene S. Kaye**, *Kelly J. Clemens*[$]*, *Daniel Hermens***

[*]Centre for Research on Ageing, Health and Well-being, The Australian National University, Canberra, ACT, Australia, [$]Macquarie University, Sydney, NSW, Australia, [**]The Brain and Mind Research Institute, Camperdown, NSW, Australia

OUTLINE

WHAT IS METHAMPHETAMINE?

Methamphetamine is the chemical N-methyl-1-phenylpropan-2-amine. Common synonyms include methylamphetamine and metamfetamine (international nonproprietary name). Methamphetamine, a powerful stimulant drug, is a controlled substance under international drug control conventions due to its highly addictive nature and serious health consequences. The effects of intoxication include euphoria, wakefulness, hypervigilance, confidence, and energy. Adverse effects include dependence, psychosis, and a range of other mental and physical sequelae, which are discussed in this chapter. On the illicit drug market,

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00070-X

methamphetamine is sold under a variety of pseudonyms such as crystal, shabu, ice, tik, ya ba, base, pure, and speed.

HISTORY

Methamphetamine was first synthesized by Nagai in Japan in 1893. It was used by military combatants during World War II and Japan experienced several subsequent epidemics of use after its release onto the broader market. A similar phenomenon was observed in Europe and the United States with amphetamine, methamphetamine's less potent chemical analog. Both drugs gained popularity as pharmaceutical agents because of their decongestant properties and their capacity to suppress appetite, alleviate depression, and overcome fatigue. Throughout the subsequent decades, many accounts of abuse of pharmaceutical preparations containing these amphetamine analogs and related incidents of psychosis were published in the scientific literature. This led to the withdrawal of some pharmaceutical products from the market and their eventual international control under the 1971 United Nations Convention on Psychotropic Substances. While methamphetamine is still available as a pharmaceutical drug in some countries (e.g. under the trade name Dexosyn® in the United States), its medical application has been largely superseded by less potent and less addictive amphetamine analogs. Today, methamphetamine is more typically associated with illicit drug use.

GLOBAL TRENDS IN USE

In 2010 the United Nations Office on Drugs and Crime estimated that there were between 14 and 53 million methamphetamine consumers worldwide. Many countries do not have accurate estimates of the prevalence of methamphetamine use. Among those that do, the past year use of the drug varies between 0 and 5% of the adult population (these figures often include the use of methamphetamine and/or amphetamine). In a broad sense, methamphetamine is a major drug problem for several countries in Southeast and East Asia, North America, and the Pacific (Australia and New Zealand). There are only small pockets of methamphetamine use within Europe (the Czech Republic and Slovak Republic); amphetamine is more common in other parts of the region. Methamphetamine has recently become a concern in South Africa, and there is evidence of clandestine manufacture emerging in South Asia, Southwest Asia, and South America, but its use in these regions is not well established.

EFFECTS OF METHAMPHETAMINE

The principle effect of methamphetamine desired by its users is the euphoria, or the more general sense of well-being, that it instates. This effect is coupled with an increase in confidence, energy and enthusiasm, alertness, and an increase in mental acuity. Intoxication tempers appetite and staves off sleep. The drug also increases respiration, elevating heart rate and blood pressure; it causes pupillary dilation and decreases micturition. Studies have found that intoxication (at least mild intoxication) speeds reaction time and can improve performance of activities that require sustained vigilance, giving its users the capacity to remain engaged in even the most tedious of tasks. However, improvements in performance are offset by changes in attentional processing and shifts in goal-directed activities (e.g. users sometimes engaging in unproductive activities for long periods of time). The drug's capacity to instate bravado can also affect judgment and lead people to engage in high-risk situations. Signs of intoxication are difficult to detect at low levels, with the drug having no obvious deleterious effects on motor coordination; the only tell-tale signs of intoxication are that the person may appear more confident, talkative, and enthusiastic than normal. At high doses, intoxication may be reflected in agitation, anxiety, hyperarousal, excitability, and incessant talking. A better indicator of use is the ongoing circadian disturbances (insomnia followed by hypersomnolence), weight loss, and the labile mood that tend to accompany regular stimulant use.

DEPENDENCE ON METHAMPHETAMINE

Methamphetamine is a highly addictive drug that has a defined dependence syndrome. Dependence on methamphetamine is characterized by tolerance, affective and physical withdrawal symptoms, users taking the drug to relieve withdrawal, and rapid reinstatement of use following periods of abstinence. Tolerance to methamphetamine is reflected in the use of higher doses, more frequent use, and a shift to more efficient routes of administration (e.g. injecting). The withdrawal syndrome from methamphetamine involves an acute phase that lasts about 1 week. Dominant symptoms include fatigue, craving, irritability, depression, anxiety, circadian disturbances, poor concentration, and anhedonia. Low-level symptoms of withdrawal can persist for a further 2 weeks after this acute withdrawal phase. The management of methamphetamine withdrawal is usually symptomatic. Although a number of targeted pharmacotherapies have undergone trials, none have been found to be effective.

NEUROCHEMISTRY

Methamphetamine belongs to a family of drugs often referred to as amphetamine-type stimulants. This class of drugs shares a ring-substituted amine structure and has similar effects on monoamine transmission. Methamphetamine increases monoamine levels via several mechanisms: methamphetamine reduces the reuptake of monoamines into the presynaptic nerve terminal by binding to, and reversing, the monoamine transporters; it causes the leakage of monoamines from the synaptic vesicles inside the presynaptic nerve terminal, increasing the amount of monoamines released when the nerve fires, and it slows the metabolism of monoamines by inhibiting monoamine oxidase. The primary action of methamphetamine is on dopamine; increases in dopamine in the brain's mesolimbocortical pathway are believed to mediate the pleasurable effects of the drug.

PHARMACOKINETICS

The bioavailability of methamphetamine depends on how the drug is taken. Methamphetamine can be ingested, injected intravenously, absorbed through the nasal mucosa when snorted, or the vapors of the heated drug can be inhaled and absorbed through the lungs (i.e. smoked). Methamphetamine injection and smoking yield high bioavailability (around 90%) and the effects of the drug peak within 10–20 min after administration. Intranasal use also yields good bioavailability (79%) with peak effects within a similar time frame. The effects from oral administration are slower to occur (45 min to 1.5 h) and bioavailability is comparatively lower at 67%. The subjective effects of methamphetamine are strongest within the first hour and diminish from 4 to 6 h after using the drug. The half-life of methamphetamine in humans is around 11 h, irrespective of the route of administration. Much of the dose is excreted unchanged in urine (37–45% of the nominal dose) alongside smaller quantities of its metabolites, which include amphetamine. Elimination time varies with the pH of the urine, this being faster when urine is acidic. Methamphetamine can usually be detected in blood or urine for 2–3 days after use. Detection in saliva is less reliable.

ISOMERS

Methamphetamine occurs in dextrorotatory and levorotatory stereoisomers (i.e. D-methamphetamine and L-methamphetamine). The D and L isomers of methamphetamine are absorbed similarly in the central nervous system, but the D isomer produces a stronger and longer-lasting high, and more intense cardiovascular effects, while the L isomer has a slightly longer half-life (approximately 13–15 h vs 10 h for D-methamphetamine). The stronger potency of the D isomer lends itself to use in pharmaceutical preparations but it also has greater abuse potential. Illicitly produced methamphetamine may contain either isomer or both isomers depending on the precursor chemicals used in its production.

CHEMICAL PRODUCTION AND PRECURSORS

Methamphetamine is synthetically produced using a variety of chemical precursors and reagents. The most common precursor chemicals seen in clandestine manufacture are phenyl-2-propanone (P2P, also known as phenylacetone), ephedrine and, most recently, pseudoephedrine. These chemicals are diverted from their legitimate use (e.g. pseudoephedrine is used as a decongestant in cold and flu medications) into clandestine drug manufacture. Manufacture from P2P produces a racemic mixture of the D and the L stereoisomers of methamphetamine, but this method of manufacture has been largely replaced by production using pharmaceutical grade pseudoephedrine, which produces racemically pure D-methamphetamine. Additional chemical refining processes can be used to remove impurities and create very pure methamphetamine solution that will yield shards of crystalline methamphetamine.

PRECURSOR REGULATION AS A METHOD OF SUPPLY CONTROL

Preventing the diversion of precursor chemicals into clandestine drug manufacture has become a keystone in efforts to combat illicit methamphetamine supply. In 1988, the United Nations instigated the Convention Against Illicit Traffic in Narcotic Drugs and Psychotropic Substances, which deemed criminal the supply and trafficking of such chemicals with the intent to manufacture methamphetamine. Various countries have additionally implemented domestic regulations to prevent the diversion of precursor chemicals from their legitimate use, including restrictions on the availability of cold and flu medications containing pseudoephedrine. Some of these regulatory changes have been associated with reductions in methamphetamine-related harms, but their efficacy is, to some extent, undermined by the capacity of clandestine chemists to develop alternative methods of manufacture and to source precursor chemicals from countries where domestic controls are lax.

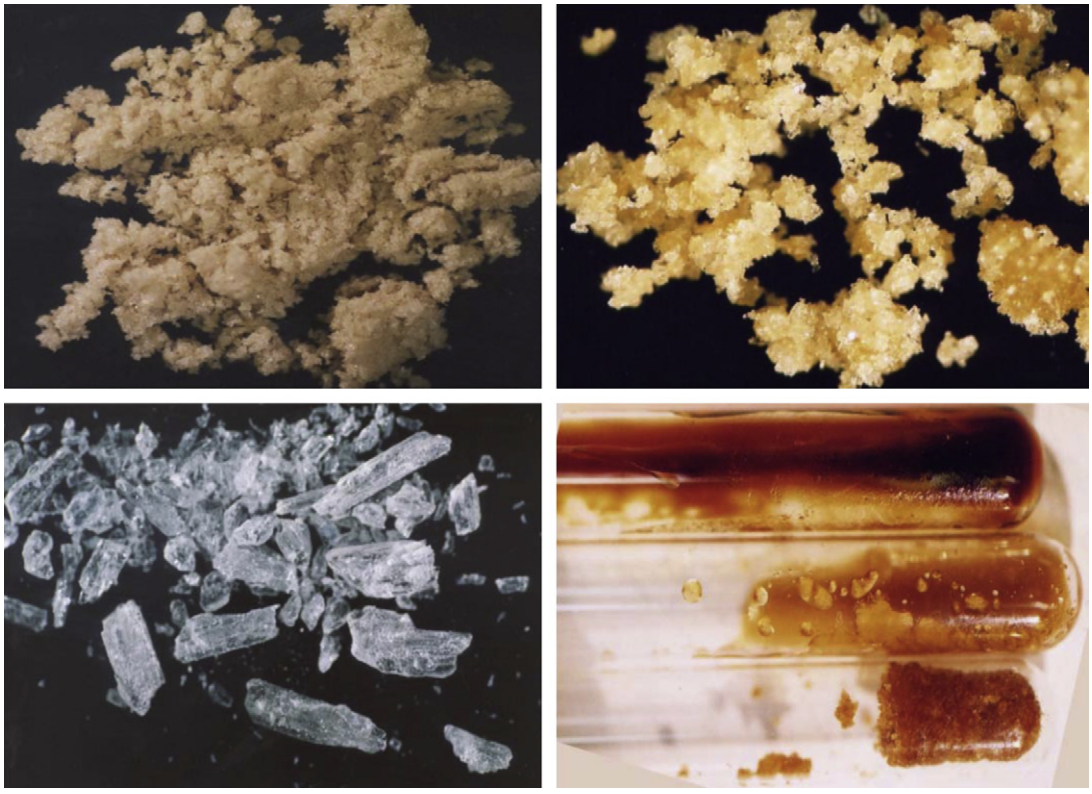


FIGURE 70.1 Methamphetamine seized in Victoria, Australia. Images provided by the Victoria Police Forensic Services Department, Chemical Drugs Intelligence Team.

ADULTERANTS

Methamphetamine that is bought and sold on the illicit drug market is typically not pure. Adulteration is necessary to control the strength of the dose being delivered; only 20–50 mg of methamphetamine is needed to induce a robust drug effect. Adulteration may also be used by drug dealers in an attempt to increase their profit margins. Contrary to popular belief, most of the adulterants used to prepare methamphetamine for sale on the illicit drug market are relatively innocuous (e.g. sucrose). Harmful constituents usually occur because of incorrect manufacturing procedures or contamination with reagents used in the manufacturing process, which can include toxic heavy metals and carcinogens.

FORMS OF METHAMPHETAMINE

Methamphetamine is an unusual drug in that it is sold on the illicit drug market in various physical forms (Fig. 70.1). Crystalline methamphetamine is probably the most widely recognized form of methamphetamine, but the drug is also sold in powder, liquid, or tablet form. In some parts of the world (especially Southeast Asia) methamphetamine is combined with ketamine, and sometimes other synthetic stimulant drugs, and sold in pills, often called ya ba (meaning crazy medicine in Thai). In Australia methamphetamine is a common ingredient in pills sold as ecstasy. This trend of marketing methamphetamine in such a variety of forms, and under different street names, has caused confusion about what drugs people are actually using and has made it difficult to monitor trends in methamphetamine use.

NATURAL HISTORY OF METHAMPHETAMINE USE

The onset of methamphetamine use typically occurs in late adolescence, after the onset of cannabis and alcohol use, with problems from these various drugs tending to coalesce in early adulthood. In line with this, most methamphetamine use is seen among young adults and consists of infrequent noninjecting use of the drug. Most people cease use as they mature into their later adulthood; however, a proportion will become dependent on the drug and may continue to use for many years.

The progression to regular use and dependence typically occurs within a few years of initiation, with

epidemiological studies suggesting that around one in five users will make this transition. Similar to the dependence on other drugs, the risk of becoming dependent on methamphetamine is influenced by a range of individual, environmental, and societal factors. More frequent drug use, and efficient routes of administration, particularly those that provide a more immediate drug effect, can also facilitate the development of dependence. While methamphetamine use overall tends to be more prevalent among males than females, males are more likely to progress to chronic dependent use.

METHAMPHETAMINE USE PATTERNS

Most methamphetamine use occurs among adolescents and young adults who use the drug in social settings (e.g. at dance venues) because it enhances confidence and increases energy. In this context, use is typically infrequent (e.g. monthly or less often) and involves noninjecting routes of administration (i.e. snorting, swallowing, or smoking the drug). Use patterns tend to become less social and more frequent with the progression to dependence. Dependent users may take the drug several times a day, and a characteristic binge and crash pattern is often observed, where users may take the drug for days on end until they eventually succumb to the effects of sleep deprivation. Occupational use is also commonly associated with methamphetamine use. The drug's ability to enhance wakefulness and improve vigilance makes it a popular drug for use among long-distance transport workers and shift workers.

METHAMPHETAMINE INJECTION

Methamphetamine available on the illicit drug market readily dissolves in water, which makes it suitable for injection. Methamphetamine use by injection tends to be associated with more problems than oral or intranasal use. Besides the risk of contracting bloodborne viruses (e.g. human immunodeficiency virus (HIV) and hepatitis B and C), and other injection-related problems (e.g. thrombosis, endocarditis), injecting users face an increased risk of dependence and related problems (e.g. financial difficulties, social problems, and poor mental health). Injecting users tend to be older, and they often have a long history of treatment attempts and involvement with the criminal justice system. It is not uncommon for methamphetamine injectors to have a history of opioid injection and heavy polydrug use. This can further exacerbate health problems among methamphetamine injectors and complicate treatment provision.

METHAMPHETAMINE SMOKING

Methamphetamine vaporizes at high temperatures, allowing the drug to be smoked. Smoked methamphetamine is absorbed in the lungs from where it directly enters the pulmonary blood supply giving a rapid drug effect that rivals injection. This efficient absorption of methamphetamine makes smoking a popular route of administration with recreational methamphetamine users and dependent users alike. Smoking can be associated with compulsive use; it conveys a greater risk of dependence than other noninjecting routes of administration (e.g. snorting or swallowing the drug) and, similarly, tends to be associated with more problems (Table 70.1).

Patterns of smoking methamphetamine are also different from those associated with other routes of administration. Crystalline methamphetamine (which is the form of methamphetamine most commonly smoked) vaporizes when heated and recrystallizes when cooled, allowing unused methamphetamine to be saved and reused later. This property of methamphetamine allows smokers to titrate their dose, inhaling small amounts of the drug at regular intervals to reinstate the high. It is also a convenient property for social users of the drug who wish to share their methamphetamine among friends. Crystalline methamphetamine is typically smoked in a purpose-made glass pipe. While crystalline methamphetamine is the form of methamphetamine most commonly smoked, other forms of the drug can also be smoked; for example, in Southeast Asia, the smoking of methamphetamine pills is common.

TABLE 70.1 Experience of a Methamphetamine Smoker

The following quote is from a 22-year-old male methamphetamine smoker who was interviewed in Sydney, Australia, in 2004. He had first tried methamphetamine when he was 15 years old and by the time he was 20 years old, he was engaging in episodic periods of heavy use. At the time of the interview he was working full-time as a manager, living in a share-house with several other people who were smoking methamphetamine, over half of his social circle was also using the drug, and he was experiencing paranoid ideation and symptoms of methamphetamine dependence

"I was on ice daily for at least two months … pretty much everyone in the house was smoking ice, dealing ice, getting high on their own supply… it's a really more-ish drug that you want to keep smoking and smoking. Even though you are feeling the effects of the ice you still feel that you've got to have that pipe and smoke it continuously… so I fell into that pattern, again, of smoking it everyday for a month. And, again, psychologically and emotionally, I had, like, not a breakdown, but I was getting there. And even though I was still going to work and everything, it was sort of like smoking ice then smoking pot to come down again, coming up again with ice, so I just became a shell of myself, and just depressed and moody and angry. I knew it was a problem… I think ice had a hold on me… and I didn't want to admit to myself that I was pretty much addicted"

METHAMPHETAMINE PSYCHOSIS

Methamphetamine can induce a transient psychotic state that is characterized by hallucinations and persecutory ideation. Less common manifestations of methamphetamine psychosis include stereotyped repetitive behavior (punding), disorganized speech, and illogical tangential thoughts. Symptoms typically last only hours and, even in severe cases, usually abate within a week of withdrawal from the drug. A minority of users experience a more prolonged symptom pattern that resembles schizophrenia, although it is often argued that this protracted form of methamphetamine psychosis would be more accurately described as a precipitation of schizophrenia in vulnerable individuals.

Methamphetamine psychosis is most often observed among heavily dependent methamphetamine users. A higher risk is associated with a younger onset of methamphetamine use, more frequent use, and a higher blood concentration of methamphetamine. The development of the more prolonged form of methamphetamine psychosis appears to have a strong familial component, whereby the presence of psychotic illness in first-degree relatives conveys a significant risk. As such, a number of candidate genes have been proposed to mediate vulnerability to (or protection against) methamphetamine's psychomimetic properties. Methamphetamine can also precipitate and exacerbate psychotic symptoms in other psychotic disorders (e.g. schizophrenia and bipolar affective disorder) and impede the effectiveness of antipsychotic medication.

Although studies consistently show elevated levels of psychosis among methamphetamine users, prevalence estimates vary widely depending on the level of methamphetamine use and the background risk for psychosis in the population being studied. Elevated levels of psychosis among methamphetamine users are not only due to the drug but also an overrepresentation of people with other psychotic disorders (e.g. schizophrenia) among methamphetamine users. Other factors associated with heavy methamphetamine use, such as sleep deprivation and a history of trauma, can further increase vulnerability to psychosis. While these background risk factors for psychosis contribute to the high prevalence of psychotic symptoms among methamphetamine users, they cannot completely account for it. Experimental studies indicate that high doses of methamphetamine can induce a psychotic reaction and this relationship has been repeatedly borne out in epidemiological studies and clinical case reports (for a description see Bell, 1973).

The symptom profile of methamphetamine psychosis is extremely similar to acute paranoid schizophrenia but the two conditions can usually be distinguished on the basis of symptom duration; methamphetamine psychosis symptoms are comparatively transient and co-occur with methamphetamine use. There is also less evidence of a residual syndrome of negative symptoms associated with methamphetamine psychosis. Knowledge about the most appropriate treatment for methamphetamine psychosis is limited, but antipsychotic drugs do alleviate the symptoms, and often sedation is sufficient because symptoms usually recede naturally with detoxification.

AGGRESSION AND VIOLENT BEHAVIOR

Historically, both methamphetamine and its less potent analog amphetamine have been associated with bizarre acts of violence and even homicide. These incidents are believed to arise from a confluence of the drug's pharmacological properties, its ability to incite paranoia, and a range of personality and contextual factors.

From a pharmacological perspective, high doses of methamphetamine and chronic exposure to methamphetamine can increase aggression. The hyperaroused state associated with intoxication can also augment aggression that is incited by a threatening situation. This relationship appears to be specific to high doses of methamphetamine, with little evidence that low doses increase aggression.

Methamphetamine-induced paranoia also plays an important role in precipitating acts of violence. Paranoid ideation can lead methamphetamine users to misinterpret events and perceive an unrealistic level of threat in their environment, accounting for the seemingly irrational and sometimes extreme nature of methamphetamine-related violence. Such situations are further exacerbated by the background risk for violent behavior among methamphetamine users, including antisocial tendencies and impulsivity, which tend to co-occur with heavy illicit drug use.

A further important consideration in understanding the relationship between methamphetamine use and violence is that illicit drug users can also get drawn into violent crimes because they are associated with a criminal underworld or because they need to fund an expensive illicit drug habit. These types of factors increase the risk of violence irrespective of the pharmacological effects of the drug.

DEPRESSION AND OTHER MENTAL HEALTH PROBLEMS

Depression, anxiety, and other mental health problems tend to be overrepresented among people who regularly use methamphetamine, as is the case with

drug users more generally. These types of mental health problems often precede drug use and may increase the risk of someone becoming dependent on methamphetamine (or other drugs). Ongoing methamphetamine use can further exacerbate some of these mental health problems, most notably psychosis (as discussed earlier), but also depression. Withdrawal from heavy methamphetamine use can induce a pseudodepressive state that lasts several days to weeks, and which is underpinned by a downregulation of monoamine activity in brain regions implicated in the regulation of mood and affect. Such methamphetamine-related mood disturbances can be sufficiently severe as to constitute a clinically meaningful entity in their own right.

NEUROTOXICITY

Chronic exposure to methamphetamine can cause a downregulation of monoamine activity and high doses can be neurotoxic. Methamphetamine neurotoxicity primarily involves the degeneration of dopamine neuron terminals, although other neurochemical systems (e.g. serotonin and noradrenaline) can also be compromised after very high levels of methamphetamine intake. Neuronal damage is largely restricted to brain regions that receive dense innervation from dopamine neurons, including the striatum and prefrontal cortex. Methamphetamine-induced neural degeneration is mediated by oxidative stress, which occurs when normal metabolic processes are unable to cope with the large efflux of dopamine following methamphetamine intake. Recent studies have implicated a range of additional neuropathological processes that may further contribute to methamphetamine-related neurotoxicity, including disruption of metabolic processes as well as gliosis and reactive astrogliosis. Together these processes result in a loss of normal neuronal function and signal transmission.

Evidence of neurotoxicity has been seen in chronic methamphetamine users, including reductions in the density of dopamine transporters, serotonin transporters, dopamine D2 receptors and vesicular monoamine transporters, and changes in metabolic activity in dopaminergic regions. These neurochemical abnormalities correlate with psychiatric symptoms and longer-duration methamphetamine use. There is evidence in humans and animals that deficits partially recover with ongoing abstinence although recovery is not always seen.

COGNITIVE DEFICITS

There are a number of studies showing poor cognitive functioning among methamphetamine users compared to non–drug-using controls. The most consistently observed deficits are in the domains of executive function and memory, with research also showing evidence of slowed information processing, poor attention, and motor problems. Cognitive deficits in methamphetamine users have been correlated with various deficiencies in the frontal and striatal regions of the brain (i.e. those brain regions affected by neurotoxicity) but these deficits are not always correlated with exposure to methamphetamine use (i.e. duration and extent of use), undermining the argument that they are due to neurotoxicity. Importantly, a host of other factors associated with illicit methamphetamine use, besides neurotoxicity, may have a detrimental impact on cognitive functioning among methamphetamine users. These include circadian disturbances and mood disturbances related to methamphetamine withdrawal, infections such as HIV and hepatitis C, head injury, and polydrug use (including brain damage from hypoxia or hyperthermia experienced during drug overdose).

HIV RISK

Methamphetamine intoxication increases libido, and some users take the drug specifically to enhance sex. Consequently, methamphetamine has been suggested as a vector for HIV transmission. The association between methamphetamine use and sexual risk behavior has been consistently observed in a range of population subgroups, including high-risk groups for HIV infection (e.g. men who have sex with men, youth in areas of high HIV prevalence). Prospective cohort studies have observed that methamphetamine use is an independent risk factor for HIV seroconversion. In addition to the relationship between sexual risk behavior and methamphetamine use, it is thought that the drug may increase susceptibility to HIV infection through immunopathological processes.

Methamphetamine use can also increase the risk of HIV and other blood-borne viruses (e.g. hepatitis B and C), through injecting drug use. Most methamphetamine use is noninjection, but injecting is sufficiently prevalent to warrant concern about its impact on HIV epidemics and blood-borne virus transmission more generally. Substantial populations of injecting methamphetamine users have been noted in Australia, New Zealand, Japan, North America, and parts of Central and Eastern Europe.

CRIME

Crime is not especially common among recreational methamphetamine users, tending to occur mainly

among heavier users of the drug. Most heavy methamphetamine users are not only funding a methamphetamine habit, but many are also heavy cannabis users, smoke tobacco, and drink alcohol and/or use other drugs socially. Couple the high cost of such regular drug use with the low income typical of chronic methamphetamine users, and it is not hard to see why many become involved in crime. Dealing drugs is by far the most common way that heavy methamphetamine users support their drug habit, although a substantial proportion also engage in theft (particularly shoplifting to obtain food and other everyday commodities) and a minority resort to more serious crimes (e.g. robbery, fraud). There appears to be a particular association between methamphetamine use and violent crime, which is likely to be related to the pharmacological effects of the drug (see section on Aggression and violent behavior). The actual prevalence of violent crimes among methamphetamine users, however, is not greatly elevated compared to that seen with other drugs (e.g. heroin), and such crimes are far less common than either selling drugs or petty theft. Crime is also most likely to be seen among younger male methamphetamine users, particularly those with a history of conduct problems and antisocial behavior — these being generic risk factors for criminal involvement which are elevated among heavy drug users.

DENTAL PROBLEMS

Excessive methamphetamine use is associated with jaw clenching, bruxism, and cracked teeth. The often cited meth mouth syndrome refers to the proliferation of dental caries and the occurrence of gingivitis in some users of the drug. Evidence suggests that this is largely caused by poor dental hygiene, resulting from a chaotic drug-using lifestyle, rather than any direct effect of methamphetamine.

CARDIOVASCULAR EFFECTS

Methamphetamine use has been associated with a range of acute and chronic effects on the cardiovascular system. Chest pain, hypertension, tachycardia, and other cardiac arrhythmias are the most commonly observed acute cardiovascular effects of methamphetamine. Other, less common, acute complications include myocardial infarction, coronary vasospasm, acute aortic dissection, and sudden cardiac death. In addition to acute forms of cardiac pathology, methamphetamine use has also been associated with chronic cardiac pathology, which, in turn, may increase the risk of an acute cardiac event.

The forms of chronic cardiovascular disease most commonly associated with methamphetamine use are coronary artery disease and cardiomyopathy. Specifically, the use of methamphetamine, chronic use in particular, is thought to contribute to the premature and accelerated development of coronary artery disease and induce dilated cardiomyopathy. Underlying coronary artery disease, cardiomyopathy, and generalized cardiomegaly (enlarged heart) have been observed among emergency department patients presenting with acute coronary symptoms following methamphetamine use, as well as among methamphetamine-related fatalities at autopsy.

The cardiovascular effects of methamphetamine are thought to be mediated primarily by the release of catecholaminergic neurotransmitters (i.e. noradrenaline and dopamine) in the peripheral nervous system, which modulates heart rate and blood pressure. High catecholamine levels are known to be cardiotoxic, causing vasoconstriction, vasospasm, tachycardia, and hypertension. While tachycardia and hypertension are associated with increased demand for oxygen to the heart muscle, vasoconstriction and vasospasm decrease the cardiac oxygen supply. The net effect of these co-occurring conditions is a deficit in the oxygen available to the heart, which can in turn cause necrosis, fibrosis, and an increase in the size of heart muscle cells (hypertrophy). Methamphetamine has also been shown to have cardiotoxic effects that are independent of the catecholamine-mediated effects described above, which are referred to as direct cardiotoxic effects, although the mechanisms underlying these effects are still unclear.

MORTALITY

Methamphetamine is less commonly associated with mortality than opioid use but deaths can result from the cardiotoxic effects of the drug, as described above, and also from an increased risk of cerebrovascular accidents. Other causes of premature death seen among methamphetamine users include suicide and accidents. A recent systematic review of mortality rates among methamphetamine (and amphetamine) users, conducted by Singleton and colleagues, found that crude mortality rates varied between 0 and 3 per 100 person years, with substantial geographic variation in mortality. It was not clear to what extent background mortality contributed to these rates, and consequently to what extent methamphetamine elevated the risk of mortality. Only one study reported a standardized mortality, which was based on the recorded mortality among a large cohort of hospitalized methamphetamine users in the Czech Republic. This study found a large elevation in deaths among methamphetamine users

compared to the expected mortality rate (standardized mortality ratio of 6.2). As pointed out by the authors, while mortality appears elevated among methamphetamine users, there is currently insufficient understanding of the factors contributing to this elevation to draw firm conclusions about the risk of mortality associated with methamphetamine.

TREATMENT AND MANAGEMENT

Few controlled trials have been conducted to evaluate the benefit of specific treatment options for methamphetamine dependence. So far none of the treatment options that have undergone trials have proved to be particularly effective, and those that have shown promise are difficult to implement on a large scale. Behavioral interventions, such as cognitive behavioral therapy and contingency management, have improved outcomes in selected trials, but a recent meta-analysis Colfax and colleagues found no conclusive evidence that behavioral interventions, overall, yield better outcomes than passive or minimal treatment interventions. Improvements were seen with high-intensity interventions, namely those that were specifically tailored to the target population and those that used adjunctive contingency management. These high-intensity interventions have not been implemented widely within community-based drug treatment services.

There is a strong demand for pharmacotherapies for methamphetamine dependence (akin to opioid pharmacotherapies). In an attempt to meet this need, a range of pharmaceuticals have been explored, both to medicate withdrawal and as potential maintenance drugs. Most of these medications have not improved outcomes enough to warrant pursuing large-scale trials or community-based implementation, and some have had adverse side effects, or carried a high risk of diversion, which has rendered them undesirable.

In the meantime, most methamphetamine users receive treatment from generic drug treatment clinics that service the needs of people with various types of drug and alcohol problems. These services include, but are not restricted to, the medical management of withdrawal symptoms, outpatient counseling services (which may include structured psychological interventions), and residential rehabilitation programs (which may incorporate various aspects of withdrawal management and counseling). These treatment-as-usual options appear to provide at least temporary relief for the majority of methamphetamine users seeking help.

A handful of generic drug treatment services have also been tailored to cater to the specific needs of methamphetamine users. In the United States, the MATRIX model of treatment was modified for stimulant use, with this program involving a structured 16-week outpatient psychosocial treatment with follow-up care. Evaluations found that the MATRIX treatment out performed treatment-as-usual for methamphetamine dependence during treatment delivery, but once people left the treatment their drug use was similar to that seen following standard care. A similar approach has been adopted in Australia with the development of outpatient stimulant treatment programs. The self-help group Narcotics Anonymous has also been adapted for methamphetamine users, this program being called Crystal Meth Anonymous, although it has not been evaluated.

One of the endeavors of providing specialized drug treatment for methamphetamine dependence, besides improving care for methamphetamine users, is to attract more methamphetamine users into treatment. Currently, treatment coverage for methamphetamine dependence is quite low. It has been estimated that less than one-third of dependent methamphetamine users have ever received treatment for their methamphetamine use. This problem is likely to reflect the lack of low-threshold interventions, such as pharmacotherapies. But it also reflects that many users feel a low need for treatment, either because they are not particularly motivated to reduce their drug use, or because they prefer to self-manage their use rather than attend a formal drug treatment program. This problem has led clinicians and researchers to embark on a creative journey of designing treatment approaches that are more likely to engage methamphetamine users in a therapeutic process. Some of these approaches are embodied within the specialized treatment options described above; others include using internet interventions and information resources to engage with methamphetamine users who are not seeking treatment, and a stepped care approach in delivering treatment, which allows the focus and intensity of the treatment intervention to be tailored to the needs of the methamphetamine user (e.g. focusing on emotional problems associated with drug use).

SUMMARY

Methamphetamine is a powerful synthetic stimulant drug that has been around for just over a century. It was initially used as a pharmaceutical agent, but its use resulted in serious adverse effects, including psychosis and dependence, leading to its control under international drug control conventions in 1971. Methamphetamine remains available as a pharmaceutical drug in some countries, but its pharmaceutical use has been largely superseded by safer amphetamine-type analogs. Today methamphetamine remains widely used as an illegal drug, with the number of current users estimated to be between 14 and 53 million worldwide. The regions

that have been most heavily affected by methamphetamine use are North America, Southeast and East Asia, Australia, and New Zealand; however, illegal manufacture and use is spreading to affect more regions, including Africa, South and Southwest Asia, and South America. While most methamphetamine use consists of infrequent recreational use among young adults, around one in five users experience dependence on the drug, resulting in chronic use and a myriad of health and social problems. Mental health is particularly compromised, with the drug having the capacity to cause a paranoid psychosis and aggressive behavior. The cardiovascular effects are also concerning, although less well recognized. Despite the risk of dependence, and the myriad of consequent health and social problems, there is a paucity of effective specialized treatment approaches to manage either dependence or the psychiatric sequelae associated with the drug's use.

List of Abbreviations

D2 dopamine 2 (receptor).
HIV human immunodeficiency virus.
P2P phenyl-2-propanone.

Further Reading

Allen, A., Cantrell, T.S., 1989. Synthetic reductions in clandestine amphetamine and methamphetamine laboratories – a review. Forensic Science International 42, 183–199.
Bell, D.S., 1973. The experimental reproduction of amphetamine psychosis. Archives of General Psychiatry 29, 35–40.
Colfax, G., Santos, G.-M., Chu, P., et al., 2010. Amphetamine-group substances and HIV. The Lancet 376, 458–474.
Cruickshank, C.C., Dyer, K.R., 2009. A review of the clinical pharmacology of methamphetamine. Addiction 104, 1085–1099.
Curran, C., Byrappa, N., McBride, A., 2004. Stimulant psychosis: systematic review. British Journal of Psychiatry 185, 196–204.
Darke, S., Kaye, S., McKetin, R., Duflou, J., 2008. Major physical and psychological harms of methamphetamine use. Drug and Alcohol Review 27, 253–262.
Dawe, S., Davis, P., Lapworth, K., McKetin, R., 2009. Mechanisms underlying aggressive and hostile behavior in amphetamine users. Current Opinion in Psychiatry 22, 269–273.
Ellinwood, E.H., 1971. Assault and homicide associated with amphetamine abuse. American Journal of Psychiatry 127, 90–95.
Kaye, S., McKetin, R., Duflou, J., Darke, S., 2007. Methamphetamine and cardiovascular pathology: a review of the evidence. Addiction 102, 1204–1211.
Krasnova, I.N., Cadet, J.L., 2009. Methamphetamine toxicity and messengers of death. Brain Research Reviews 60, 379–407.
Marshall, B.D.L., Werb, D., 2010. Health outcomes associated with methamphetamine use among young people: a systematic review. Addiction 105, 991–1002.
Mendelson, J., Uemura, N., Harris, D., et al., 2006. Human pharmacology of the methamphetamine stereoisomers. Clinical Pharmacology and Therapeutics 80, 403–420.
Scott, J.C., Woods, S.P., Matt, G.E., et al., 2007. Neurocognitive effects of methamphetamine: a critical review and meta-analysis. Neuropsychology Review 17, 275–297.
Singleton, J., Degenhardt, L., Hall, W., Zabransky, T., et al., 2009. Mortality among amphetamine users: a systematic review of cohort studies. Drug and Alcohol Dependence 105, 1–8.
United Nations Office on Drugs and Crime, 2010. World Drug Report 2010. United Nations, New York.
Vocci, F.J., Appel, N.M., 2007. Approaches to the development of medications for the treatment of methamphetamine dependence. Addiction 102 (Suppl. 1), 96–106.

Relevant Websites

http://www2.cochrane.org/reviews/ – Cochrane review library for evidence on methamphetamine-related interventions.
http://www.crystalmeth.org – Crystal Meth Anonymous.
http://www.matrixinstitute.org – Outpatient addiction treatment centers.
http://www.meth.org.au/ – Self-help for methamphetamine users.
http://ndarc.med.unsw.edu.au/ – Information and education resources from the National Drug and Alcohol Research Centre, Australia.
http://www.nida.nih.gov/infofacts/methamphetamine.html – National Institute on Drug Abuse facts on methamphetamine.
http://www.stvincents.com.au – St Vincent's Hospital, Sydney, has an outpatient stimulant treatment program.
http://www.methinformation.org/ – UCLA Integrated Substance Abuse Programs (ISAP) information on methamphetamine treatment.
http://www.unodc.org/ – United Nations International Drug Control Conventions and global methamphetamine trends.

CHAPTER

71

Hallucinogens

*Silvia S. Martins**, *Grace P. Lee**, *Zila M. Sanchez*[$], *Paul Harrell**, *Lilian A. Ghandour***, *Carla L. Storr*[§]

[*]Columbia University, NY, USA [$]Universidade Federal de São Paulo, CEBRID (Brazilian Center of Information of Psychotropic Drugs), São Paulo, SP, Brazil [**]American University of Beirut, Riad El-Solh, Beirut, Lebanon [§]University of Maryland School of Nursing, Baltimore, MD, USA

OUTLINE

Hallucinogens are a diverse group of substances that cause perceptual and cognitive distortions; they can evoke mood and thought changes as well as out-of-body experiences. Traditionally derived from plants or based on plant-derived compounds, their use in medicinal and religious ceremonies goes back for centuries. Many designer drugs and two other drug classes, dissociative and deliriant drugs, are also often classified under this category. Other club drugs that overlap with hallucinogens, such as gamma-hydroxybutyrate, are not included in this chapter. This chapter provides a brief overview of the neurobiology and mechanism of action of hallucinogens, including atypical hallucinogens. The physiologic and psychological effects of hallucinogens are described, followed by discussion on the epidemiology of hallucinogen use, diagnosis of hallucinogen intoxication and abuse/dependence, comorbidity of hallucinogen use with other drug-use and psychiatric

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00071-1

disorders, prevention and intervention strategies to reduce hallucinogen use, and treatment of hallucinogen-related disorders.

NEUROBIOLOGY OF HALLUCINOGENS

Classification of Hallucinogen-Type Drugs

Almost all hallucinogens contain nitrogen and are classified as alkaloids. One possible way of classifying the hallucinogens is by their chemical structure and the receptors they act on. There are two main groups: (1) the indolealkylamines, which are further divided into substituted tryptamines (e.g. psilocybin (*N*,*N*-dimethyl-4-phosphoryloxytryptamine), psilocin from *Psilocybe* mushrooms, DMT (5-methoxy-dimethyltryptamine) from ayahuasca, *Psychotria viridis*) and lysergic acid derivatives (D-lysergic acid diethylamide (LSD) from the ergot fungus); and (2) phenylalkylamines, structurally related to catecholamine neurotransmitters and represented by drugs such as mescaline (3,4,5-trimethoxy-phenethylamine) and mescaline-substituted structures (2,5-dimethoxy-4-methylamphetamine (DOM) and 2,5-dimethoxy-4-bromoamphetamine (DOB)). Figure 71.1 shows the chemical structure of the hallucinogens that are considered prototypical (typical hallucinogens: LSD, psilocin/psilocybin, and mescaline).

There are also three other groups of drugs that may have hallucinogenic effects but are not always classified as hallucinogens (or typical hallucinogens): (1) designer drugs such as 3,4-methylenedioxymethamphetamine, also known as ecstasy (MDMA), 3,4-methylenedioxyamphetamine (MDA), 3,4-methylenedioxy-ethylamphetamine (MDE), and *para*-methoxymethamphetamine; (2) dissociative anesthetics (such as phencyclidine (PCP) and ketamine), which are arylcycloalkylamines, and (3) cannabinoids. Cannabinoids are presented in another chapter of this book. Figure 71.2 shows the chemical structure of the hallucinogens that are considered atypical.

Although the action of hallucinogens is not yet fully elucidated, it is known that typical hallucinogens act on serotonergic pathways.

All of these drugs are serotonergic (5-HT_2) receptor agonists or partial agonists. The 5-HT_2 receptors represent a family of receptors that are composed of three subpopulations: 5-HT_{1A}, 5-HT_{2B}, and 5-HT_{2C}. The cortical and limbic regions are the areas with the main concentration of this subtype of serotonergic receptors. The effects of each hallucinogen are not identical because there are also other serotonergic receptors involved with different affinities for each one. While phenylalkylamine hallucinogens are somewhat selective in their affinity to 5-HT_{2A} receptors, the indolalkylamines are relatively nonselective for 5-HT receptors,

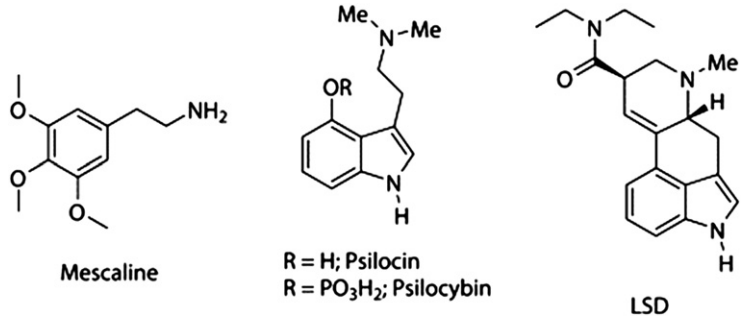


FIGURE 71.1 Chemical structures of prototypical hallucinogens: mescaline, psilocybin and psilocin, and LSD.

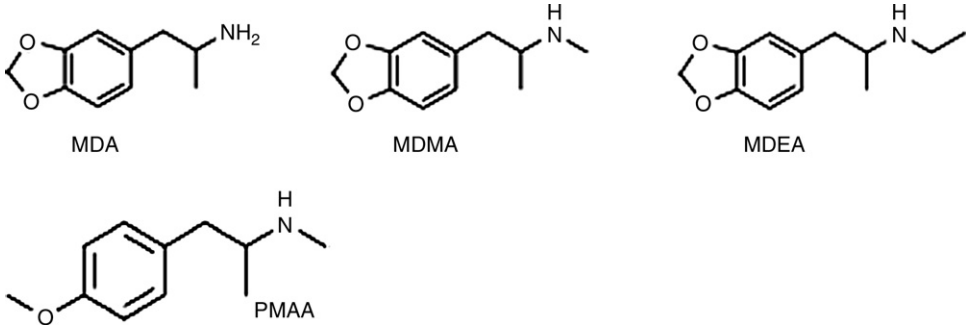


FIGURE 71.2 Chemical structure of atypical hallucinogens.

displaying moderate to high affinity for diverse 5-HT_1 and 5-HT_2 subtypes. This happens mainly because this group is very similar in chemical structure to serotonin and fits into different serotonin receptors. In addition to the psychotomimetic effects of the serotonergic pathways, it is also hypothesized that 5-HT_{2A} receptors play a major role in the discriminative stimulus actions of these agents. However, not only do the discriminative stimulus properties of the drug vary depending on the experimental conditions and context, but they may also depend on the dose of drug used.

Mechanism of Action of Atypical Hallucinogens

Designer Drugs

The synthetic drugs that have been pharmaceutically created as mind-altering drugs, such as MDA, MDMA, MDE, and many others, are not always classified as typical hallucinogens but have a common amphetamine-type mechanism of action and are sometimes classified as hallucinogenic amphetamines. Several studies in rodents and humans have shown that MDMA increases central extracellular serotonin levels. Even though MDMA initially induces increased extracellular levels of serotonin, multiple exposures and/or large doses lead to reductions in brain serotonin levels. MDMA has a low affinity for dopamine receptors, which suggests that all the psychomotor and psychological dopamine-mediated effects of MDMA are a consequence of the dopamine release induced by the drug, increasing the extracellular concentration of dopamine in the nucleus accumbens. There is sufficient evidence showing that the effect of this group of drugs on the serotonergic system is closely associated with activation of dopaminergic pathways and release of dopamine. The enhanced release of dopamine by these drugs may partially result from the reversion of dopamine transport.

Dissociative Anesthetics

Many of the effects of PCP and other arylcyclohexylamines are mediated by dopamine. PCP and ketamine are antagonists of *N*-methyl-D-aspartic acid (NMDA) (which blocks NMDA receptors on dopaminergic neurons in the tegmental ventral area) instead of having a direct effect on dopaminergic neurons. The hyperactivity that PCP induces in rodents has been proved to be related to midbrain dopamine areas. NMDA receptors are the primary mediators of glutamate-induced stimulation of midbrain dopaminergic neurons and the glutamatergic routes of substantia nigra from the prefrontal cortex is a major determinant of dopaminergic activity levels.

Because atypical hallucinogens act directly or indirectly to enhance the concentration of dopamine in the nucleus accumbens area, they may induce dependence. On the other hand, as indolalkylamines do not act in the reward area of the brain, they do not induce dependence.

PHYSIOLOGICAL EFFECTS

Hallucinogens are generally considered to be physiologically safe molecules because, although they can produce altered states of consciousness, they do so at doses that are not toxic to mammalian organ systems. Researchers believe that hallucinogens do not cause damaging effects on cardiovascular, renal, or hepatic functions due to their lack of affinity for the biological receptors and targets that mediate vital vegetative functions. Although hallucinogens may have few long-term physiologic effects, they have been shown to produce some acute effects such as dizziness, weakness, tremors, nausea, drowsiness, paresthesia, and blurred vision. LSD use has also been shown to correspond to signs of sympathetic arousal such as increased pulse and blood pressure, dilated pupils, piloerection, hyperreflexia, and slight fever. Another study examining the potential dose effects of psilocybin found that electrocardiogram and body temperature were not affected by psilocybin dosage, but plasma concentrations of thyroid-stimulating hormone, prolactin, cortisol, and adrenocorticotropic hormone did increase at the peak of the effects of high doses of psilocybin. However, the plasma levels of the hormones decreased over time, thus further supporting the low risk of long-term physiologic effects of hallucinogens.

MDMA can cause dry mouth, nausea, loss of appetite and sweating up to 1 week after ingestion. It is also well established that MDMA use can cause hyperthermic syndromes, which might lead to death.

PSYCHOLOGICAL EFFECTS

In contrast to other drugs that act on the central nervous system, where the action is usually predictable, the effects of hallucinogens are heavily dependent on the expectations of the user and the environment in which the use takes place. Furthermore, a user's response to repeated administration of the same drug and dose could also vary in that high doses do not always produce effects that are similar to the effects of low doses but at greater intensity. Typical psychiatric effects of hallucinogens observed in clinical settings include alterations in mood, tension, distorted time sense, difficulty in expressing thoughts, depersonalization, dreamlike feelings, and visual hallucinations (Hollister, 1984). LSD can also induce psychoses. In case reports of LSD

psychosis, the most common symptoms were found to be mood swings, visual hallucinations, mania, grandiosity, and religiosity, suggesting that LSD psychotics are fundamentally similar to schizophrenics in genealogy, phenomenology, and course of illness. Ayahuasca tea is a hallucinogenic beverage often used in religious and ritual settings throughout the western Amazon Basin. Barbosa and colleagues found that among subjects who consumed ayahuasca tea for the first time in a ritual experience, many experienced extraordinary visual experiences, a prominent sense of inner calm and harmony, a sense of a superior and powerful presence, elucidating thoughts, and alterations in self-body image. Similarly, in a sample of Australian adults, DMT use was reported to have been a personally meaningful or insightful experience that produced feelings of euphoria and intense visual hallucinations. Another study that pooled raw data from eight double-blind placebo-controlled experimental studies to create a sample of 100 healthy subjects who received one to four oral doses of psilocybin found that although there was a dose effect in changes in mood, perception, thought, and self-experience, most subjects described the experience as pleasurable, enriching, and nonthreatening. There has been no evidence of long-term psychological, cognitive, and behavioral deficits of hallucinogen use. In addition, in contrast to many other drugs of abuse, hallucinogens have low dependence qualities and are not generally considered to be reinforcing substances. It has been suggested that most drugs with dependence liability affect dopaminergic (DA) transmission. However, most hallucinogens lack affinity for either the DA receptors or DA uptake transporters, and as a result do not activate the brain reward pathways as observed for other drugs with high abuse and dependence potential such as opiates and cocaine.

MDMA can cause acute side effects including difficulty concentrating, anxiety, depressed mood, dissociative feelings, and insomnia. Long-term ecstasy use might lead to neurotoxic consequences. Furthermore, tablets purchased on the street often contain dangerous additives and/or other illicit drugs.

EPIDEMIOLOGY OF HALLUCINOGEN USE

United States

According to the National Survey of Drug Use and Health (NSDUH), in 2009, 14.8% of the US population aged 12 years and older had already used a hallucinogen in their lifetime (9.4% LSD, 2.5% PCP, and 5.7% MDMA), 1.8% had used hallucinogens in the past year (0.3% LSD and 1.1% MDMA), and 0.5% were past-month users (0.1% LSD and 0.3% MDMA). Past-month hallucinogen use was highest among adolescents aged 12–17 years (overall 0.9%: 0.1% LSD, 0.1% PCP, and 0.5%MDMA) and young adults aged 18–25 years (overall 1.8%: 0.3% LSD and 1.1% PCP). Only a small proportion of respondents said they had used hallucinogens as their first illegal drug (2.1%). The average age of first PCP, LSD, and MDMA use in 2009 was 16.8, 18.2, and 20.2 years, respectively. The number of new hallucinogen initiates aged 12 years and older in 2009 (1.3 million) was higher than the number of new initiates in 2003 (886 000).

In addition, other surveillance surveys have focused specifically on US adolescent behaviors and drug use. The 2009 Youth Risk Behavior Surveillance Study, which obtained data from a nationally representative sample of students in all grades, estimated that 8% of the youth interviewed had already used hallucinogens in their lifetime. Hallucinogen users were more likely to be male than female (10.2% versus 5.5%), and more likely to be White (9.0%) or Hispanic (7.9%) versus Black (3.3%). With regard to MDMA use, 6.7% of the high school students had already used MDMA in their lifetime and were more likely to be male than female (7.6% versus 5.5%), and more likely to be Hispanic (8.2%) than White (6.4%) or Black (5.1%). Another well-known school-based study, Monitoring the Future (MTF), surveys students of selected grades (8th, 10th, and 12th) and follows them up to young adulthood. MTF findings indicate that LSD is no longer the most widely used hallucinogen. Both the NSDUH and the MTF studies show declines in LSD and MDMA lifetime and past-year use among adolescents from 2002 to 2009, but increases in LSD (MTF and NSDUH) and MDMA (NSDUH only) past-month use among youth aged 12–17 years. Past-year MDMA use declined by about 50% among youth from 2002 to 2005 according to both surveys. NSDUH data indicated an increase in past-year MDMA use between 2005 and 2009 among youth, but MTF data showed no significant change in MDMA use over this same period. Both surveys found that the prevalence of LSD use has remained unchanged among young adults in recent years, but MDMA use appears to be on the increase again.

Data from the US Drug Abuse Warning Network on emergency room visits show that the number of drug-related emergency department (ED) visits involving MDMA increased significantly from 10 220 visits in 2004 to 22 816 visits in 2009, representing an 81% increase; most of these visits were made by patients aged 18–29 years. Almost 80% of these visits involved MDMA in combination with alcohol and other illegal drugs. The number of drug-related ED visits involving PCP increased from 31 342 in 2004 to 36 719 in 2009, representing a 17% increase, and those involving LSD increased from 2146 in 2004 to 4028 in 2009, representing an 88% increase.

Australia, Canada, and Europe

In Australia, using data from the 2007 National Drug Strategy Household Survey, it was estimated that 8.9% of the general population reported ever using MDMA, with 3.5% reporting past-year use; these estimates have increased over the past 12 years. On the other hand, data from the Australian School Students Alcohol and Drug Survey estimated that lifetime and past-year MDMA use among high school students remained relatively stable between 1996 and 2005 at around 4% and 3%, respectively. In Canada, with the exception of marijuana, hallucinogens were the drugs that respondents to the Canadian Addiction Survey (2004) most commonly reported using in their lifetime (11.4%). In Europe, using data from the 2007 European School Survey Project on Alcohol and Other Drugs study, it was estimated that 2–3% of all high school students aged 15–16 years in the 35 participating countries had already used hallucinogens (including LSD and MDMA) in their lifetime. The highest European rates of hallucinogen use were found to be 5% in the United Kingdom, Ireland, and the Czech Republic, and 4% in Moscow. The overall consumption levels of LSD and hallucinogenic mushrooms among European adults are generally low and have been largely stable in recent years.

Of the approximately 2.5 million (0.8 %) European adults who used MDMA in the last year, virtually all are in the 15–34 years age group.

DIAGNOSIS OF HALLUCINOGEN INTOXICATION AND HALLUCINOGEN ABUSE AND DEPENDENCE

Hallucinogen intoxication is diagnosed after recent use of one or more hallucinogens followed by marked anxiety or depression, ideas of reference, fear of losing one's mind, paranoid ideation, impaired judgment, or impaired social or occupational functioning that occurred while using the drug or immediately after using it. In addition, the individual experiences subjective intensification of perceptions, depersonalization, derealization, illusions, hallucinations, and synesthesias while using the drug or immediately after using it; and might experience clinical symptoms such as tachycardia, pupillary dilation, sweating, palpitations, blurring of vision, tremors, and incoordination.

According to *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV) criteria, hallucinogen abuse is diagnosed in an individual in the same way that abuse of other substances is diagnosed. That is, a hallucinogen user with at least one of four symptoms representing recurrent negative psychosocial consequences of use or hazardous use is diagnosed with hallucinogen abuse, as long as the individual has never met criteria for hallucinogen dependence. On the other hand, hallucinogen dependence is diagnosed when an individual meets, within a 12-month period, at least three of six criteria related to physical dependence, impaired control over substance-use behavior, and increased salience of consumption (the withdrawal criteria do not apply to hallucinogens, see DSM-IV-TR). For DSM-V, there are recommendations to combine hallucinogen abuse and dependence into a new unidimensional disorder: hallucinogen use disorder.

Few studies have focused on MDMA abuse and/or dependence. In a study conducted by Cottler and colleagues in the United States, nearly half (43%) of the adolescents who used MDMA more than four times were found to meet the diagnostic criteria for ecstasy dependence, as shown by continued use despite knowledge of physical or psychological harm, withdrawal effects, and tolerance (or diminished response); and 34% met the criteria for MDMA abuse. In a study of 200 Taiwanese adolescent MDMA users from a juvenile abstinence center, 22% of the participants met the adopted DSM criteria for ecstasy dependence, with 24% meeting the criterion for tolerance and 6.5% meeting the criterion for withdrawal. On the other hand, data from the German Early Developmental Stages of Psychopathology (EDSP) study found the prevalence of DSM-IV ecstasy use disorders to be fairly low (1.6%).

COMORBIDITY WITH OTHER DRUG USE, DRUG-USE DISORDERS, AND PSYCHIATRIC DISORDERS

Dual Diagnosis with Another Psychiatric Morbidity

The psychiatric consequences of hallucinogen use, particularly MDMA (ecstasy), remain understudied, particularly with regard to the long-term effects. Enduring outcomes, particularly in human populations, have not been fully established. Most of the literature delineating anxiety and social behavioral problems in the long term has been based on animal models. Moreover, epidemiological research studies have also mostly focused on targeted populations (rave party goers). Nonetheless, most (not all) of the available evidence based on animal models and human populations, generated using cross-sectional, longitudinal, and laboratory studies, suggests moderate to strong associations between hallucinogen use (particularly ecstasy) and psychiatric disorders.

In a cross-sectional study on a nationally representative adult population (18 years and older) from the United States (2001–2002), Keyes and her colleagues

found statistically significant strong associations between past 12-month MDMA use and axis I disorders including current anxiety disorders, panic disorders, and specific phobia; current mood disorders were only linked to former (not current) ecstasy use, potentially indicating a longer-term effect of use. This study's findings are in agreement with those from laboratory studies that have shown psychological impairment and depressive symptoms up to 2.5 years after cessation of MDMA use. The German study also corroborates the possible link to major depression and other mental health disorders, whereby MDMA users were significantly more likely to be diagnosed with a DSM-IV defined major depressive disorder after 4 years, in addition to panic attacks, phobia, and generalized anxiety disorder, compared with nonusers, and even other illegal drug users; former and current ecstasy users also exhibited higher odds of any personality disorder, specifically antisocial and paranoid personality disorders, as well as higher odds of histrionic and schizoid personality disorders (in current users only) compared with users of nonillegal drugs. Findings of a study using the 2001–2002 National Study of Alcohol and Related Conditions data conducted by Compton and his colleagues illustrated that hallucinogen use disorders (both abuse and dependence) were related to higher likelihood of adult antisocial personality disorder in both men and women (stronger in the latter). Although understudied, the possible link with suicidality is quite worrisome; one study by Kelly et al. found an increased risk for suicide attempts among males with hallucinogen use disorders.

Co-occurrence with Alcohol and Other Illegal Drug Use

Most of the literature has focused particularly on the co-occurrence of MDMA and other substances, and has shown that hallucinogen use, particularly MDMA, rarely occurs in the absence of alcohol and other illegal drugs. In a study by Keyes and her colleagues, an alarming 44% of adult past-year ecstasy users were reported to also use more than three classes of illegal drugs in addition to alcohol, not to mention their higher odds of binge drinking (62% of current users) and alcohol abuse. Epidemiological studies on young hallucinogen users, MDMA in particular, also find a strong association with cigarette, alcohol, as well as other illegal drug use (whether they were used to intensify the MDMA experience or minimize its adverse after effects). Polydrug use may be even higher among MDMA users (versus any other hallucinogen user). One recent study on youth aged 12–17 years found a higher prevalence of past-year use of alcohol, tobacco (and nicotine dependence), opioids, cocaine (and cocaine use disorders), and tranquilizers compared with all other hallucinogen users. Very few studies have examined hallucinogen use disorders (abuse, dependence, and diagnostic orphans). One study by Wu and colleagues (2009) pointed to an increased odds of marijuana and alcohol use disorders among adolescents diagnosed with hallucinogen dependence or abuse. It has been shown that alcohol–hallucinogen use is a very common alcohol–drug combination in the United States, whether in a sample of 176 adolescents or 43 000 adults. Other studies have found a threefold increase in marijuana use disorder among those with hallucinogen dependence (in addition to a fourfold increased odds of a major depressive episode). Ecstasy users were found to be at a sixfold increased odds of alcohol use disorders (compared with nonusers) and approximately a twofold increase compared with other illicit drug users. Thus, hallucinogen users in general, and MDMA users in particular, are polydrug users, often using more than one illegal substance during the same time period.

Current Issues and Future Considerations

Most studies have either examined hallucinogens in general, or MDMA in particular, and very few have examined substance use and the psychiatric correlates of other hallucinogens. Most recently, a study examining *Salvia divinorum* (a potent naturally occurring hallucinogen that is used infrequently by the general population) found that recent and former users had a greater odds of past-year depression and alcohol- or drug-use disorders, compared with past-year alcohol and drug users who did not use *S. divinorum*. Future studies on hallucinogen use need to focus, to the extent possible, on the potential effects of particular substances, exclusive of ecstasy, in order to begin to understand and delineate the potential mechanisms for intervention, be it at the prevention and/or treatment level.

The high comorbidity with other substances and mental health disorders limits our ability to tease apart the effects of MDMA from that of other substances or psychiatric diagnoses. This opens up the discussion of whether MDMA is independently related to other psychiatric disorders, irrespective of its co-occurrence with other substance use. The issue of polydrug use by ecstasy users and the difficulty in delineating independent effects have been raised by several researchers. One study found no relation between past-year and lifetime ecstasy use and any of the anxiety or depression measures once the frequency of other drug use was controlled for, and another did not establish any significant differences between ecstasy users and controls with regard to the prevalence of mental disorders.

There is a possible bidirectional nature of MDMA use with other illegal substance use. A study by Martins and

colleagues on youth aged 12–21 years using 2002–2003 NSDUH data revealed that use of marijuana, cocaine, and heroin predicted subsequent MDMA use, and earlier MDMA use was also predictive of later illegal drug-use initiation (although for marijuana, the association was stronger for the pathway from early marijuana use to later ecstasy initiation; the opposite was true for cocaine and heroin). The EDSP study from Germany also found that alcohol, nicotine, and other drugs predominantly followed MDMA use, and that very few ecstasy users had initiated use prior to the onset of these other substance-use disorders. Thus, hallucinogen use, particularly MDMA, may be a marker for earlier substance-use or mental health problems, and a predictor of subsequent drug-use and mental disorders. Now that most studies have established a correlation between hallucinogen use and other substance-use and psychiatric disorders, future studies must focus on adopting the longitudinal approach needed to address the issue of temporality and begin understanding causal effects.

This section summarizes the current knowledge on hallucinogens and their comorbid conditions, but it is important to also highlight some of the limitations inherent in all of the studies that have generated the above-mentioned conclusions and affected their validity, as acknowledged by the authors. Besides the cross-sectional nature of most surveys, and their sampling method, issues such as ensuring proper classification of users (e.g. absence of biological markers), sample size (e.g. power calculations), choice of controls (e.g. other drug use or nondrug use), and adequate observation time (e.g. long enough to miss acute effects and short enough to observe longer-term consequences of use/abstinence) are some of the methodological considerations that future studies must address. Moreover, the purity, intensity, frequency, or quantity of MDMA use is often not measured, and future national surveys must begin to measure the differential effects and perhaps a dose–response relationship that various amounts of hallucinogen use may have on the risk of substance-use and other psychiatric outcomes (or vice versa). Also, it would be interesting to distinguish between polydrug use within the same year on various different occasions and coingestion, which has severe implications on the users' well-being as well as treatment interventions.

PREVENTION/INTERVENTION

Cognitive Behavioral Skills

School-based approaches to substance abuse prevention emphasizing cognitive behavioral skills training to enhance social resistance skills and other general life skills have been found to produce both short-term and long-term effects on the prevention of hallucinogen use. Botvin and colleagues (2000) were one of the first to examine the effects of such programs on hallucinogen use using a random sample of 3597 seventh graders from rural and suburban schools in New York State further randomized into either an intervention or control group. Throughout the 15-session program, students in the intervention group were taught cognitive behavior skills for building self-esteem, resisting advertising pressure, managing anxiety, communicating effectively, developing personal relationships, and asserting rights. Furthermore, information regarding the immediate negative consequences of drug use, the decreasing social desirability of use, and the actual prevalence rates among adults and adolescents was provided. Booster sessions designed to review and reinforce the material covered during the first year of intervention were provided in eighth and ninth grades. Conversely, students in the control group received no information on cognitive behavioral skills and drug use. A subsample of 447 students from the original cohort was reinterviewed at the end of 12th grade (mean age 18.1 years), and 13% of the participants in the intervention group reported lifetime use of hallucinogens compared with 21% of the control group. Such findings suggest that prevention programs with cognitive behavioral skills training carried out during middle school produced observable and durable prevention efforts after high school with regard to hallucinogen use.

Project Towards No Drug Abuse (Project TND) is a drug abuse prevention program targeting high-risk adolescents attending continuation high schools, which are alternative schools that youths attend after they transfer out of the regular school system due to functional problems such as conduct problems and drug use. Project TND utilized a classroom-based curriculum that included sessions teaching effective listening skills, chemical dependency issues, alternative coping skills, and encouraged making no-drug-use decisions. In addition to the classroom curriculum, a community program component was added to allow students to participate in activities. To examine the long-term effects of the program, Sun and colleagues (2006) selected a subsample ($n = 725$) of the original cohort of 3813 students was recruited from 21 CHSs with complete annual follow-up up to 5 years after the program. Of the 725 participants, 32% were randomized to the control group, 33% were in the classroom curriculum only group, and 34% were in the classroom and community group. At the 1-year follow-up assessment, 30-day hard drug (hallucinogen, cocaine, stimulants, inhalants, depressants, PCP, steroids, and heroin) use was higher in the control group than both the classroom-only and classroom and community groups. Similarly, by the 5-year follow-up, the

frequency of hard drug use in the last month by those in the classroom-only intervention was less than half that in the control group; for those in the classroom and community intervention it was one-fifth that of the control group (Sun et al., 2006). Such findings demonstrate that Project TND has long-term effects on prevention of drug use in continuation high school students.

Assertive training aims to enable individuals to do what they really want to do in social situations. As a substance-use prevention strategy, assertive training rests on the assumption that many youths who would otherwise abstain from using drugs reluctantly participate in drug use because they lack the interpersonal skills necessary to separate themselves from such situations (Horan et al., 1982). Horan and colleagues conducted an assertive training program among 72 eighth graders to assess its effect on drug use. A behavioral measure of assertiveness was first taken among the 142 ninth graders enrolled in a public school, and the least assertive 36 males and 36 females were selected to participate in the program. Individuals were then randomized into assertion training treatment, placebo treatment, and no-treatment groups. The students in the assertion training group received five 45-min sessions with counselors instructing them about assertiveness and live modeling of assertive responses to a particular training stimulus involving peer pressure to use drugs through role playing. The placebo treatment involved discussions with students about assertiveness, peer pressure, and drug use; the no-treatment control students received no instructions from counselors and did not participate in any discussions. Immediately after the intervention, assertiveness was found to be significantly higher in the assertion training group than both the placebo and no-treatment groups. At the 3-year follow-up, although no significant difference was found in the mean frequency of lifetime hallucinogen use among the three groups, the assertion training group did report significantly lower frequency in lifetime hard drug use (which included hallucinogens, barbiturates, amphetamines, cocaine, and heroin) than the placebo and no-treatment groups. Horan and colleagues hypothesized that the lack of significant differences found in hallucinogen use among the treatment groups was due to the low prevalence (6.9%) of hallucinogen use in the sample. However, the lower usage of hard drugs, which included hallucinogens, in the assertion training group could indicate that assertion training could be a productive strategy for hallucinogen use prevention.

Knowledge and Expectancies

Studies have demonstrated that expectancies, particularly positive expectancies, are strongly related to substance use. In addition, knowledge about substance effects and past use of substances also shape expectancies. The In the Know preventive intervention program created by Stickle and colleagues aimed to increase knowledge about drugs and alcohol and to change the expectancies about their effects among 385 students in grades 6 to 12 from urban public schools. Students were randomized into the alcohol, tobacco, marijuana intervention group ($n = 123$), the alcohol, tobacco, marijuana control group ($n = 62$), the hallucinogen intervention group ($n = 136$), and the hallucinogen control group ($n = 64$). Only findings from the hallucinogen intervention and control groups are presented here. Students in the intervention group were shown a 15-min videotape that discussed and illustrated the effects of hallucinogens, followed by a discussion of the tape led by their health teachers. Students within the control groups were not shown the videotape and did not participate in a discussion with their health teacher. All students completed pretest and posttest survey questionnaires. The hallucinogen intervention group had significantly higher posttest scores on knowledge than the hallucinogen control group. In addition, the intervention group showed a significant increase in knowledge scores and a significant decrease in positive expectancies for hallucinogen use from pre- to posttest; the knowledge scores and expectancies for the control group did not change significantly over time. The results of the study suggest that the In the Know substance abuse prevention curriculum was effective in changing youths' knowledge and expectancies of drug use, which could in turn decrease their risks for later substance use.

Attitude and Behaviors

Due to the generally low prevalence of hallucinogen use in the general population, many prevention programs combine hallucinogens with other classes of drugs with low prevalence, such as cocaine, opiates, and inhalants, to create a hard drug outcome category. The Positive Youth Development Collaborative is one such program that incorporated its curriculum targeting substance-use attitudes and behaviors among urban minority youths into after-school programs. In this program, a total of 304 adolescents, mean age 14.5 years, 76% African Americans, and 53% males, were selected from five after-school programs, and were then randomized into either the intervention ($n = 149$) or control ($n = 155$) groups. Those in the intervention group were taught effective decision-making skills, particularly in substance use, with emphasis on identifying positive personal attributes and setting goals for healthy living. Control group participants participated in the usual activities provided by their after-school program. Seven months after enrollment, Tebes and colleagues found that those in the intervention group were significantly

more likely to view drugs as harmful at the end of the program than control participants. Although the 30-day incidence of hard drug use (hallucinogens, amphetamines, cocaine, heroin, nonprescription methadone, tranquilizers, or inhalants) in the intervention group increased over time from baseline (10.8%) to the 1-year follow-up (12.8%), the increase was significantly lower than the increase found within the control group (from 7.8% to 16.8%). Such reductions in the progression of substance use among adolescents could protect them against later increased or escalating use.

TREATMENT

Various disorders are associated with hallucinogen use and, as with most disorders, health care providers have found some success with treatment, but also experienced some frustration.

Classic Hallucinogens (LSD, Psilocybin, Mescaline)

Abuse and Dependence

Dependence on LSD, psilocybin, and mescaline is rare, perhaps partially due to a quickly developed, but short-lived, tolerance that renders chronic use ineffective. In addition, no withdrawal syndrome has been identified. However, hallucinogen users are often dependent on other substances. Thus, treatment methods common to most substances of abuse, such as 12-step programs or relapse prevention, may be useful here as well. However, there appears to be no formal research on the effectiveness of these methods in reducing use of hallucinogens.

Hallucinogen Intoxication/Psychosis

The abuse of hallucinogens, particularly LSD, psilocybin, or mescaline, may lead to experiences of panic or psychosis, colloquially referred to as a bad trip. These experiences can involve distressing sensory, somatic, and psychological effects, and have even resulted in suicides. Interpersonal support can be helpful, including gently holding a hand, arm, or shoulder, although, in extreme cases, an individual may react violently to touch. Conveying a sense of security, calm, and empathy is preferable to attempting to distract the user from the effects of the drug. It may also be useful to remind the user that the effects are temporary. Prolonged psychotic reactions have been reported but are rare.

Hallucinogen-induced panic can be treated pharmacologically with oral benzodiazepines. For example, 20 mg of diazepam is reported to end dysphoric anxiety within 30 min. Diazepam is recommended over lorazepam due its more rapid onset. The literature is more divided on the treatment of psychotic reactions. Some suggest against using antipsychotics, because the effects may be abrupt, unpleasant, intense, and result in subsequent psychological problems. Haloperidol is sometimes recommended for psychotic reactions, but some research suggests haloperidol can exacerbate symptoms. Risperidone has also been recommended, but, again, caution may be warranted. Olanzapine or ketanserin (a 5-HT_{2A} agonist not approved for use in the United States) may be a more useful antipsychotic. In cases of acute hypertension, intravenous labetalol can be used. Due to the effectiveness of these treatments, psychiatric emergency interventions in response to hallucinogen use are rarer now than in the past and, due to the potential for escalation of an adverse reaction, should be avoided if possible. In the rare cases of prolonged psychosis due to LSD usage, case studies have reported some success with first generation neuroleptics, electroconvulsive therapy (ECT), lithium, and the serotonin precursor, 5-hydroxytryptophan.

Hallucinogen Persisting Perception Disorder (HPPD)

According to DSM-IV, HPPD is characterized by persistent re-experiencing of various perceptual symptoms that were experienced during hallucinogen intoxication. The symptoms need to be persistent for a diagnosis of HPPD, as opposed to the transient experiences referred to as flashbacks. The disorder appears to be rare, given that millions of individuals have consumed hallucinogens; relatively few cases of HPPD have been reported. Accordingly, most treatment studies of HPPD are either case studies or involve very small sample sizes (e.g. 2–16 participants), which precludes rigorous designs. Treatments with some reported success include psychotherapy, behavior modification, the use of sunglasses, and at least 12 different pharmacological agents including clonazepam, benzodiazepines, clonidine, and a combination of fluoxetine and olanzapine. Worsening of HPPD has reportedly occurred due to exposure to phenothiazines, risperidone, and serotonin-selective reuptake inhibitors (SSRIs).

MDMA

Abuse and Dependence

Similar to the classic hallucinogens, MDMA is rarely taken more than once or twice a week due to quickly developing, but rapidly decaying, tolerance. In addition, users often report needing at least 1 day to recover after use. MDMA dependence has been reported, but the only longitudinal study of MDMA abusers of which we are aware found that over 90% of those classified as

dependent were no longer dependent 3 years later, suggesting that ecstasy dependence may be an invalid diagnostic classification, at least in terms of predictive validity.

However, some individuals do use MDMA frequently at very high doses, which may lead to adverse physical or cognitive complications. In addition, most MDMA users also abuse other substances and some users seek treatment for their ecstasy use. Treatment-seeking is believed to be rarer for MDMA, however, than for other substances, such as opioids. Treatment-seeking users of MDMA and other drugs can have more severe problems, in terms of mental health and other drug use, than users of other substances. MDMA users may be more likely to present for treatment due to other mental or nonecstasy substance-use disorders. Traditional substance abuse programs, such as 12-step modalities, may be useful, although there does not appear to be any formal research on treatment for MDMA, and it has been suggested that traditional treatment services may not be attractive to people with MDMA problems.

MDMA-Associated Disorders

Although the exact cause is unclear, use of MDMA is associated with at least three mental disorders: depression, anxiety, and psychosis. Cases for which clinicians can have some confidence in attributing the disorders to the use of ecstasy are rare. As a result, treatment studies are limited to animal research, case studies, and observational research. Some experimental evidence with rats suggests that fluoxetine, and perhaps other SSRIs, may be relatively ineffective in treating MDMA-induced depression. Case studies have reported some amount of success treating patients experiencing MDMA-induced depression with sertraline, mirtazapine, and ECT. Case studies of three patients suggest that MDMA-induced panic disorder with agoraphobia, an anxiety disorder, can be treated with serotonergic antidepressant drugs. Based on an observational study of 23 patients, olanzapine appears to be effective for ecstasy-induced psychotic disorder. These promising results should be cautiously followed up with more research.

Conclusion

This chapter provides a brief overview of the neurobiology of classic and atypical hallucinogens and how they act to cause perceptual and cognitive distortions. The most common physiologic and psychological effects of this group of drugs are described. The overall lifetime consumption level of hallucinogens among adults is low, at approximately 10–15% worldwide, but, in all countries, the recent use estimates among the young are at their highest. Intoxication, abuse, and dependence can occur, but hallucinogens, different from other illegal drugs, are not highly reinforcing. As with other drug groups, comorbidity with other psychiatric disorders and use of other substances within the same time period is very common. Prevention and intervention strategies to reduce hallucinogen use and treatment of hallucinogen-related disorders are similar to those used for other drugs.

SEE ALSO

Ecstasy/MDMA, Adolescent Substance Use: Symptoms and Course, Models of Relationships between Substance Use and Mental Disorders, Epidemiology of Addiction

List of Abbreviations

ECT	electroconvulsive therapy
ED	emergency department
EDSP	Early Developmental Stages of Psychopathology
HPPD	hallucinogen persisting perception disorder
LSD	D-lysergic acid diethylamide
MDA	3,4-methylenedioxyamphetamine
MDMA	3,4-methylenedioxymethamphetamine
NMDA	*N*-methyl-D-aspartic acid
NSDUH	National Survey of Drug Use and Health

Glossary

3,4-Methylenedioxymethamphetamine (MDMA, ecstasy) an atypical hallucinogen, MDMA is a designer drug that can induce euphoria, a sense of intimacy with others, and diminished anxiety.

Ayahuasca hallucinogenic infusion, usually mixed with dimethyltryptamine-containing plants, often consumed as a beverage in religious and ritual settings throughout the western Amazon Basin.

Comorbidity the presence of one or more disorders in addition to a primary disorder.

Designer drugs atypical hallucinogens that are pharmaceutically created as mind-altering drugs; they are sometimes classified as hallucinogenic amphetamines.

Dimethyltryptamine (DMT) hallucinogen that is structurally analogous to serotonin and is the primary psychoactive in ayahuasca.

Dissociative anesthetics atypical hallucinogens that either directly or indirectly enhance dopamine concentration in the nucleus accumbens region, thereby inducing dependence.

Hallucinogens diverse group of substances that cause perceptual and cognitive distortions, and can evoke mood and thought changes as well as out of body experiences.

Indolealkylamines a group of hallucinogens that is further divided into substituted tyrptamines and lysergic acid derivatives.

Hallucinogen abuse abuse is diagnosed if a user who has never met criteria for dependence and has at least one of four DSM-IV symptoms representing recurrent negative psychosocial consequences of use or hazardous use.

Hallucinogen dependence dependence is diagnosed if a user meets at least three of six DSM-IV physical dependence symptoms, impaired control over substance-use behavior, and increased salience of consumption within a 12-month period.

Hallucinogen intoxication use of hallucinogens that leads to marked psychological symptoms such as anxiety and depression, and physiologic symptoms such as palpitations and tremors.

Hallucinogen persisting perception disorder (HPPD) a rare disorder that consists of a persistent re-experiencing of various perceptual symptoms that may occur during hallucinogen intoxication.

Lysergic acid diethylamide (LSD) semisynthetic hallucinogen that can produce psychological effects such as altered thinking processes, altered sense of time, and spiritual experiences.

MDMA-associated disorders three mental disorders (i.e. depression, anxiety, and psychosis) that are associated with MDMA use.

Mescaline naturally occurring hallucinogen that can cause cross-intolerance with some other hallucinogens such as LSD and psilocybin.

Phencyclidine (PCP, angel dust) hallucinogen that can be recreationally ingested, smoked, or snorted.

Phenylalkylamine a group of hallucinogens that are structurally related to catecholamine neurotransmitters; includes drugs such as mescaline and mescaline-substituted structures.

Psilocybin hallucinogen found in mushrooms that are used both recreationally and ritually.

Psychosis bad trips that can occur after hallucinogen intoxication, which may involve distressing sensory, somatic, and psychological effects.

Serotonergic (5-HT_2) receptors family of receptors mainly concentrated in the cortical and limbic regions to which hallucinogens bind.

Further reading

Aghajanian, G.K., Marek, G.J., 1999. Serotonin and hallucinogens. Neuropsychopharmacology 21, 16S–23S.

Botvin, G.J., Baker, E., Dusenbury, L., Botvin, E.M., Diaz, T., 1995. Long-term follow-up results of a randomized drug abuse prevention trial in a White middle-class population. Journal of the American Medical Association 273, 1106–1112.

Botvin, G.J., Griffin, K.W., Diaz, T., et al., 2000. Preventing illicit drug use in adolescents: long-term follow-up data from a randomized control trial of a school population. Addictive Behaviors 25 (5), 769–774.

Glennon, R.A., 1990. Do classical hallucinogens act as 5-HT_2 agonists or antagonists? Neuropsychopharmacology 3, 509–517.

Halpern, J.H., Sherwood, A.R., Hudson, J.I., Yurgelun-Todd, D., Pope, H.G., 2005. Psychological and cognitive effects of long-term peyote use among native Americans. Biological Psychiatry 58, 624–631.

Hollister, E., 1984. Effects of hallucinogens in humans. In: Jacobs, B.L. (Ed.), Hallucinogens: Neurochemical, Behavioral, and Clinical Perspectives. Raven Press, New York, pp. 19–33.

Levinthal, C.F., 2005. LSD and other hallucinogens. In: Levinthal, C.F. (Ed.), Drugs, Behavior, and Modern Society. Allyn and Bacon, Boston, MA, pp. 135–158.

Martins, S.S., Ghandour, L.A., Chilcoat, H.D., 2007. Pathways between ecstasy initiation and other drug use. Addictive Behavior 32, 1511–1518.

Pierce, P.A., Peroutka, S.J., 1989. Hallucinogenic drug interactions with neurotransmitter receptor binding sites in human cortex. Psychopharmacology 97, 118–122.

Simons, J.S., Dvorak, R.D., Lau-Barraco, C., 2009. Behavioral inhibition and activation systems: differences in substance use expectancy organization and activation in memory. Psychology of Addictive Behaviors 23, 315–328.

Stickle, T.R., Terranova, A.N., 2003. Program Evaluation of "In the Know" Substance Abuse Prevention Curriculum. Syndistar Education, New Orleans, LA.

Stone, A.L., Storr, C.L., Anthony, J.C., 2006. Evidence for a hallucinogen dependence syndrome developing soon after onset of hallucinogen use during adolescence. International Journal of Methods in Psychiatric Research 15 (3), 116–130.

Studerus, E., Kometer, M., Hasler, F., Vollenweider, F.X., 2010. Acute, subacute and long-term subjective effects of psilocybin in healthy humans: a pooled analysis of experimental studies. Journal of Psychopharmacology.

Weaver, M.F., Schnoll, S.H., 2008. Hallucinogens and club drugs. In: Galanter, M., Kleber, H.D. (Eds.), The American Psychiatric Publishing Textbook of Substance Abuse Treatment. American Psychiatric Publishing, Inc, Arlington, VA, pp. 191–199.

Wu, L.T., Ringwalt, C.L., Weoss, R.D., Blazer, D.G., 2009. Hallucinogen-related disorders in a national sample of adolescents: the influence of ecstasy/MDMA use. Drug and Alcohol Dependence 104 (1–2), 156–166.

Relevant Websites

NIDA InfoFacts – http://www.drugabuse.gov/publications/drugfacts/hallucinogens-lsd-peyote-psilocybin-pcp.

NIDA Research Report – http://www.drugabuse.gov/ResearchReports/Hallucinogens/Hallucinogens.html.

NIDA for Teens – http://teens.drugabuse.gov/mom/mom_hal1.php.

ONDCP – http://www.whitehousedrugpolicy.gov/drugfact/hallucinogens/index.html.

SAMHSA – http://store.samhsa.gov/facet/Substances/term/Hallucinogens?headerForList=.

The Good Drugs Guide – http://www.thegooddrugsguide.com/drug-types/types-of-hallucinogens.htm.

Treatment Center – http://www.treatment-center.com/drug-alcohol-info/hallucinogens.php.

CHAPTER

72

Ecstasy/MDMA

*Danielle E. Ramo**, *Christian Grov*[$]*, *Jeffrey T. Parsons*****

[*]University of California, San Francisco, CA, USA, [$]Brooklyn College of the City University of New York, Brooklyn, NY, USA, [**]Hunter College of the City University of New York, New York, NY, USA

OUTLINE

WHAT IS ECSTASY/MDMA

Ecstasy (3,4-methylenedioxymethamphetamine (MDMA), also called Adam, E, X, XTC eccie, hug drug, or a roll/rolling) is a synthetic, psychoactive drug that is chemically similar to the stimulant methamphetamine and the hallucinogen mescaline. Ecstasy produces feelings of euphoria, increased energy, emotional warmth, and distortions in time, perception, and tactile experiences.

Ecstasy is taken orally, usually as a capsule or tablet. Tablets, which typically contain from 50 to 150 mg of the active drug MDMA, are usually imprinted with a popular icon such as a cartoon character, animal, or corporate logos (e.g. Mitsubishi, Motorola). Users sometimes refer to the drug by these imprints (e.g. a blue dolphin pill). Pills can vary in color, white being the most common. Ecstasy is typically purchased in the setting where it is used, most commonly at bars, nightclubs, or raves, although it is increasingly used in more conventional settings such as house parties. Prices range from $20 to $40 per tablet, and it is not uncommon for tablets to be adulterated with other chemicals, including aspirin, caffeine, dextromethorphan, pseudoephedrine, amphetamine, and heroin. Although MDMA tends to be the primary substance found in ecstasy pills, less than half of tablets contain MDMA only; most tablets contain either MDMA and other substances or only substances other than MDMA. Depending on a variety of factors (including the pill's content, other adulterants, and user's metabolism), the high from ecstasy typically begins within 30 min to 1 h after consumption and lasts between 4 and 6 h. Throughout this chapter, we will use the term Ecstasy

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00072-3

to refer to pills sold as ecstasy, and use the term MDMA to describe only the specific compound if relevant.

Ecstasy can produce confusion, depression, sleep problems, drug craving, and severe anxiety (for a more detailed description, see the section on Consequences of ecstasy use). These problems can occur soon after taking the drug or occasionally even days or weeks after taking the drug. In addition, severe ecstasy users perform more poorly than nonusers on certain types of cognitive or memory tasks, although some of these effects may be due to the use of other drugs in combination with ecstasy. Research in animals indicates that MDMA can be harmful to the brain. Although similar neurotoxicity has not been shown definitively in humans, animal research on MDMA's damaging properties strongly suggests that MDMA/ecstasy is not a safe drug for human consumption. This is currently an area of active research.

HISTORY OF MDMA

MDMA was developed in Germany in the early 1900s as a parent compound to be used to synthesize other pharmaceuticals. E. Merck Pharmaceuticals filed a patent for the drug in Germany in 1912, but it remained unnoticed until a large toxicological study was done under a classified contract with the US Army after World War II. During the 1970s, in the United States, some psychiatrists began using MDMA as a psychotherapeutic tool, despite the fact that the drug had never undergone formal clinical trials or received approval from the US Food and Drug Administration (FDA) for use in humans. Nevertheless, the drug gained a small following among psychiatrists in the late 1970s and early 1980s, with some even calling it 'penicillin for the soul' because it was perceived to enhance communication in patient sessions and reportedly allowed users to achieve insights about their problems. It was also during this time that MDMA first started becoming available on the street as ecstasy. In 1985, the US Drug Enforcement Administration (DEA) banned the drug, placing it on its list of Schedule I drugs, corresponding to those substances with no proven therapeutic value.

In the 1990s there was a resurgence of ecstasy use, with use proliferating in dance clubs and rave scenes. Ecstasy became known as a club drug and was one of the most commonly used drugs of this type. Ecstasy use increased among young people throughout the 1990s in the United States, peaking in 2001. In the early 2000s, there was a widespread public safety campaign to warn young people about the dangers of ecstasy as a party drug, but that effort declined as use dropped off. In 2009, there was a resurgence of ecstasy use in the United States, with reported use in mainstream settings such as bars and homes in addition to clubs and raves.

Ecstasy was initially popular among Caucasian adolescents and young adults in the nightclub scene or raves. More recently, the profile of the typical ecstasy user has changed, with the drug now affecting a broader range of ethnic groups. Ecstasy has been particularly popular among urban gay males; some report using ecstasy along with other clubs drugs (e.g. ketamine, γ-hydroxybutyrate (GHB), crystal methamphetamine) and/or other legal and illegal substances.

NEUROBIOLOGY OF ECSTASY USE

MDMA affects the brain by increasing the activity of the monoamine neurotransmitters serotonin, dopamine, and norepinephrine. Like other amphetamines, MDMA causes an increase in neurotransmitter activity in that more of these neurotransmitters are released from their storage sites in neurons. Compared with methamphetamine, MDMA causes greater serotonin release and somewhat lesser dopamine release. The serotonin system plays an important role in regulating mood, aggression, sexual activity, sleep, and sensitivity to pain. The excess release of serotonin by MDMA likely causes the mood elevating effects experienced by ecstasy users. However, by releasing large amounts of serotonin, ecstasy causes a depletion of this neurotransmitter, likely contributing to the negative behavioral and psychological after effects that users often experience for several days after taking MDMA.

EPIDEMIOLOGY/PATTERNS OF ECSTASY USE

Use

In 2008, estimates of the current (past year) prevalence of ecstasy use among young adults in the United States and abroad range from 3 to 6%, with lifetime estimates around 10%. Rates vary somewhat by data source. The National Survey on Drug Use and Health (NSDUH), conducted by the Office of Applied Studies of the Substance Abuse and Mental Health Services Administration (SAMHSA), collects information on the prevalence, patterns, and consequences of alcohol, tobacco, and illegal drug use and abuse in the general US civilian noninstitutionalized population aged 12 years and older. In 2009, NSDUH reported that 5.7% of the US population aged 12 years and over had ever used ecstasy, 1.1% had used in the past year, and 0.3% had used in the past month. Ecstasy use in the United

States is rising; lifetime use rates rose steadily in the first decade of the 2000s, from 4.3% in 2002 to 5.7% in 2009 (Fig. 72.1). Between 2008 and 2009, estimates of the number of past month ecstasy users rose 37% from 555 000 in 2008 to 760 000 in 2009. The majority of ecstasy users are young adults, with past month use rates of 0.5% among those aged 12–17 years, 1.1% among those aged 18–25 years, and 0.1% among those aged 26 years or older. Initiation of ecstasy use began rising in 1993 when there were 168 000 new users. By 2000, there were 1.9 million new users, but that number decreased to 642 000 new users in 2003, and has steadily increased since then. In 2009 there were an estimated 1.1 million new users of ecstasy, suggesting that it has once again gained popularity.

The US Monitoring the Future Survey (MTF) is a survey conducted by the University of Michigan's Institute for Social Research for the National Institute on Drug Abuse (NIDA) and assesses licit and illicit substance use among US adolescents in grades 8, 10, and 12. It reported that past year use of ecstasy peaked in 2001, declined from 2002 to 2005 and then rebounded in 2006–2008. In 2009, trend lines leveled off in all grades. In recent years, there have been declines in perceived risk of ecstasy use in all three grades, declines in disapproval of ecstasy use among grades 8 and 10, and declines in perceived availability of ecstasy in all three grades.

Surveys of club drug users indicate that ecstasy is one of the most commonly used club drugs. After cocaine, ecstasy tends to be the most commonly reported club drug used, and up to three-quarters of young people surveyed in bars, clubs, and raves report that it was the first club drug they used. Among club drug users recruited at raves, bars, and nightclubs, median days of ecstasy use among those who use the drug was 2.5 days per month in one study. Among ecstasy users, the most commonly reported contexts of ecstasy use are raves, bars, and at home with others. Some work suggests that there are patterns of ecstasy use situations, such that some ecstasy users do not use in many situations, others use in social situations when they are feeling good, and others follow a third pattern of use in varied situations, including internal and external positive and negative states.

The Drug Abuse Warning Network (DAWN) is a public health surveillance system that documents drug-related visits to emergency departments (EDs) throughout the United States. In 2007, ecstasy use was involved (either directly or indirectly) in the fifth largest number of visits to the ED in the United States (after cocaine, heroin, marijuana, and stimulants). Ecstasy-related visits to the ED increased from 10 220 in 2004 to 16 749 in 2006, and then declined to 12 748 in 2007. Of the visits in 2007, more males (7607) visited the ED than females (5141), and those aged 21–24 years accounted for the largest numbers of visit compared with all other ages (Fig. 72.2).

Addictive Potential

For some people, ecstasy can be addictive. A survey of young adult and adolescent ecstasy users found that 43% of those who reported ecstasy use met the accepted diagnostic criteria for dependence, as shown by continued use despite knowledge of physical or psychological harm, withdrawal effects, and tolerance (or diminished response with continued use of the same amount). These results are consistent with those from similar studies in other countries that suggest a high rate of dependence among ecstasy users.

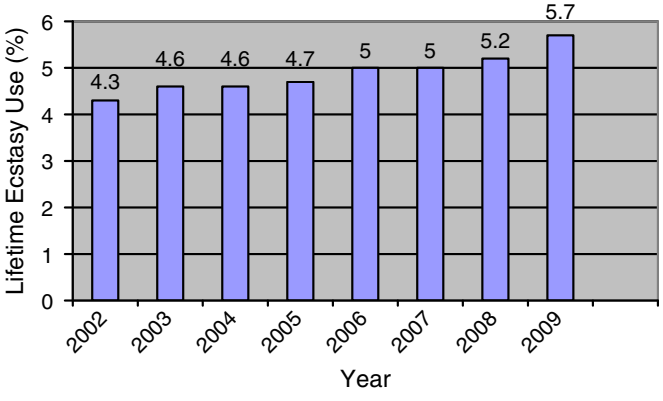


FIGURE 72.1 Lifetime ecstasy use among persons aged 12 years or older from the National Survey on Drug Use and Health: percentages, 2002–2009.

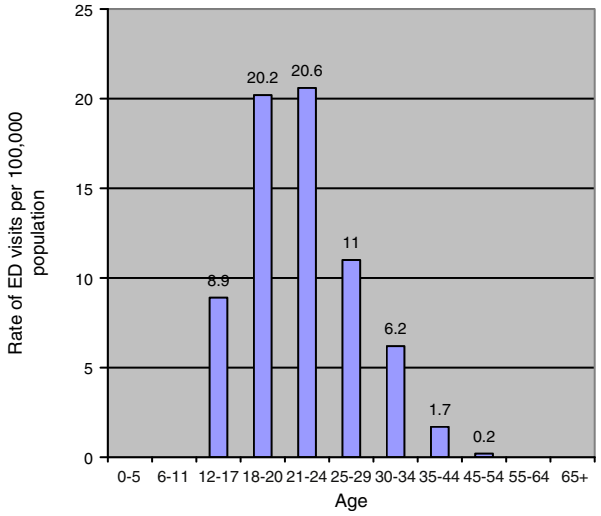


FIGURE 72.2 Rates of ED visits per 100,000 population involving ecstasy by age and gender, 2007.

Although ecstasy is currently classified as a type of hallucinogen and its withdrawal is not recognized in the *Diagnostic and Statistical Manual of Mental Disorders Fourth Edition Text Revision* (DSM-IV-TR), there is evidence for the association of withdrawal symptoms with ecstasy abstinence. Ecstasy abstinence-associated withdrawal symptoms include fatigue, loss of appetite, depressed feelings, and trouble concentrating. Findings from latent class analysis indicate that ecstasy users have a significantly higher risk of dependence than lysergic acid diethylamide (LSD) users. Of the ecstasy users who met the criteria for dependence, use despite knowledge of physical or psychological harm was the most prevalent dependence criterion (63%), while withdrawal and tolerance were also commonly reported dependence criteria. In this sample, 34% met the criteria for abuse, and hazardous use was the most commonly reported abuse symptom. Only 23% of the ecstasy-using sample did not meet the criteria for either abuse or dependence.

Research on ecstasy users compared to other hallucinogen users indicates that there are higher rates of dependence among those who use ecstasy (approximately one in five) compared to other hallucinogen users (one in six), possibly because ecstasy is more available these days and the subjective effects are less severe than some club drugs (e.g. LSD). Among ecstasy users, prevalence of hallucinogen abuse, subthreshold dependence, and dependence were 4.9, 11.9, and 3.6%, respectively. The majority with hallucinogen abuse displayed subthreshold dependence. Most with hallucinogen dependence also met the criteria for abuse. Among young adults who frequent nightclubs and other venues where ecstasy is sold, 17.4% of those who met the criteria for club drug dependence said their drug of choice was ecstasy.

Studies of ecstasy use disorders have examined which clusters of abuse and dependence symptoms best characterize ecstasy users. For example, a study examining distinctive subtypes of ecstasy users based on 24 abuse and dependence symptoms underlying standard DSM-IV-TR criteria found four latent classes representing clearly differentiated diagnostic clusters including a group of negatives, or subthreshold users endorsing few abuse and dependence symptoms (27.1%), a group of mild dependents, who had features of dependence for a select group of symptoms (34.6%), a moderately dependent group, with a similar profile of dependent symptoms as the mild dependents but also reporting some abuse symptoms (18.1%), and a severe dependent group with a distinct profile of abuse and dependence symptoms (20.1%). Studies such as those mentioned above highlight that while ecstasy is currently classified as a hallucinogen, there may be differences in the symptoms most commonly exhibited by ecstasy users compared to other hallucinogen users, and there are distinct types of ecstasy users.

ECSTASY USE AND SPECIAL POPULATIONS

Several studies have examined the sociodemographic determinants and patterns of drug use within specific population subgroups, including ethnic minorities; gay, lesbian, or bisexual males and females; juvenile detainees; injection drug users; high school or secondary school students; and college students. Since ecstasy is primarily used by young people, some work has examined its use among adolescents who are in substance abuse treatment. Lifetime rates of ecstasy are approximately 32%, and it is the second most commonly used club drug after LSD.

Although club drug use in general (including ecstasy use) has been predominantly a white, adolescent, and young adult phenomenon in the United States, club drug use among African-Americans and Hispanics is not uncommon; 45.3% of Hispanics and 56.4% of African-Americans reported a lifetime history of club drug use in one study conducted in New York City. Ecstasy was the third most commonly used club drug behind phencyclidine and LSD, with 21.7% of African-American and Hispanic youth and young adults reporting lifetime use. Of the current ecstasy users, 63.8% had used it 3 days per month or more, 31.9% used it 1–3 days per week, and 4.3% used it 4–7 days per week. Hispanics were significantly more likely than African-Americans to have used ecstasy (22.4% versus 11.7%). In a study of African-American and Hispanic heroin and cocaine users in New York City, correlates of lifetime ecstasy use included younger age, being born in the United States, and current homelessness. Compared to African-American nonintravenous drug users (non-IDUs), Hispanic non-IDUs, and white IDUs were significantly more likely to have a history of lifetime ecstasy use while African-American IDUs were significantly less likely.

Another widely studied group is men who have sex with men (MSM). There is a high rate of lifetime exposure to a variety of club drugs (including ecstasy, methamphetamine, ketamine, and GHB) in this population, with past 6-month club drug use rates in one study as high as 13.7%, and mean frequency of use of more than six times in that period. Among MSM, the use of multiple club drugs (particularly methamphetamine) over time is linked, and rates of ecstasy and other club drug use declined steadily in this population from 2002 to 2007. Among MSM club drug users, ecstasy is the most commonly used club drug, with one study showing 83% of club drug using MSM having used

ecstasy. Data are mixed about whether men identified as bisexual are more or less likely to use club drugs compared to men identified as gay. Furthermore, there is very limited data on rates of use among MSM identified as heterosexual. Among MSM, the majority of the men initiated use of club drugs with cocaine, followed by ecstasy, ketamine, methamphetamine, and GHB. Variations in patterns are related to both age and level of polydrug use of club drugs.

Various correlates of ecstasy use among MSM have been identified. Among MSM who use club drugs on a regular basis, there are high rates of a prior suicide attempt (including high rates of multiple suicide attempts), high rates of lifetime exposure to multiple types of drugs, high rates of current polydrug use (including multiple types of club drugs), and high rates of current depressive symptoms. Although high rates of condom use are reported in some types of sexual exchanges, data show multiple types of sexual risk among MSM who use club drugs regularly, including high rates of unprotected anal intercourse and concomitant other drug use among both young MSM and their sexual partners. Other correlates of ecstasy use among MSM include younger age, less education, more male sex partners, more one night stands with men, more visits to bars or clubs and sex clubs or bathhouses, likelihood of having been the victim of physical domestic violence, more gay/bisexual friends, disclosure of their sexual orientation to more friends, family members, and coworkers, and higher levels of gay community participation and affiliation. Ecstasy use has an effect on HIV status and progression of disease. Ecstasy use is associated with shorter time to seroconversion in HIV-negative MSM.

Some work has examined ecstasy use among ethnic minority MSM. For Hispanic MSM, rates of use range from 15 to 20% for recent use and are approximately 35% for lifetime use. Among Asian/Pacific Islander MSM, ecstasy is the most common drug used in conjunction with sex, and being high or buzzed on ecstasy is significantly associated with unprotected anal intercourse in this population.

PREDICTORS/CORRELATES OF ECSTASY USE

Research on sociodemographic factors associated with club drug use continues to be a main focus worldwide. Findings on sociodemographic determinants vary across studies, depending on the target populations, cultural environment, types of substances investigated, study designs, and the analytic methods used. A number of common risk factors have been identified, which can be grouped into categories.

Demographic Characteristics

According to large epidemiological studies, recent users of ecstasy tend to be young adults aged 18–21 years and residents of metropolitan areas. White young adults are more likely than others to use ecstasy, and among those who do use, whites tend to use more of the drug more often. Further, white and younger individuals are more likely to use a variety of other drugs compared to those of other ethnicities and older ecstasy users.

Males and females tend to use ecstasy at similar rates. Despite these suggestions of gender equity for rates of use, some studies have suggested that females are more likely to report negative health consequences related to the use of club drugs, specifically ecstasy. One explanation posits that this gender difference is related to the more intense subjective effects of ecstasy reported by women. Some research also suggests that females more likely than males to report recent use of multiple club drugs.

Several studies have documented differential rates of drug use by people identified as gay, lesbian, and bisexual compared with their heterosexual counterparts. In college samples, gay, lesbian, and bisexual students are more than two times as likely to have used ecstasy in the past year compared to heterosexual. In community samples, gay, lesbian, and bisexual young adults are significantly more likely to use ecstasy, cocaine, and methamphetamine over the course of their lives, as well as to be active users of club drugs, compared with their heterosexual peers. Among both male and female young adults throughout the United States, those identified as mostly hetero, bisexual, or homosexual are associated with a greater prevalence of past year ecstasy use compared to identified as completely hetero. Lesbian and bisexual women report significantly higher lifetime rates of ecstasy use compared to heterosexual women.

Other Drug Use

Use of tobacco, alcohol, and other illicit substances is associated with use of ecstasy. Using these other substances is associated with greater likelihood that youth and young adults will initiate ecstasy use. Data from the National Survey of Parents and Youth (NSPY), a longitudinal, nationally representative household survey of youth and their parents, indicates that initiation of ecstasy use is predicted by an adolescent's early initiation of smoking, drinking, or marijuana use. In particular, early initiation of either marijuana use or both smoking and drinking increases a child's risk for ecstasy use initiation. Data from the NSDUH has pointed to a gateway phenomenon whereby earlier marijuana initiation is associated with subsequent ecstasy initiation, which in turn is associated with later

cocaine and heroin initiation. While there are multiple routes in and out of substance use initiation, ecstasy initiation seems to play a role in the subsequent initiation of cocaine and heroin.

Current use of ecstasy is associated with increased likelihood of alcohol and other drug use. Ecstasy users use more types of other drugs and are more likely to be diagnosed with a substance use disorder. In a study that classified ecstasy users by the patterns of other drugs they used, it was found about 20% of ecstasy users were not likely to use other drugs, slightly more than half followed a moderate range of other substance use, and 28% were likely to follow a pattern of wide-range drug use in which they were highly likely to also use a number of other licit and illicit substances.

Using ecstasy increases the likelihood of having problems with other drugs of abuse. Relative to other drug users, those who recently use ecstasy are twice as likely to meet criteria for alcohol, marijuana, and pain reliever/opioid use disorders. They are also about twice as likely as former ecstasy users to meet criteria for marijuana, cocaine, and tranquilizer use disorders. Prescription drugs are of particular concern. In a study of the nature, extent, and consequences of prescription drug abuse among 143 ecstasy users in Miami, prescription drug abuse was reported by 87% of the sample; alprazolam (57%), oxycodone (36%), hydrocodone (32%), and diazepam (30%) were cited most often. Prescription drug abusers were more likely to report polydrug use, drug treatment histories, risky drug use behaviors, and symptoms of depression. They also reported numerous physical, psychological, and social consequences of prescription drug abuse.

Even though ecstasy users are more likely to use other drugs, they may not be more likely to use other drugs on the same days as they use ecstasy. Use of alcohol and other drugs is more likely before and during ecstasy intoxication than after. Use of ecstasy on a particular night may not be associated with any greater likelihood of using any other intoxicating drug, and use of other drugs on nights involving ecstasy use may simply reflect a natural history of drug use nights that begins with alcohol, progresses to more intoxicating drugs, and ends with little drug use.

Mental Health Problems

Mental health problems are another correlate of ecstasy use. Ecstasy users endorse more symptoms of general distress than nonusers. For example, among US high school students in the Youth Risk Behavior Survey (YRBS), disordered eating (fasting, diet product use, and vomiting or laxative use) is associated with ecstasy use. Those who use ecstasy are more likely to be diagnosed with current anxiety, specifically panic disorder, and specific phobia, and personality disorders compared with nonillicit drug users. In a study of young adult ecstasy users in Ohio, 55% had at least one lifetime disorder, with major depression (35.3%) and antisocial personality disorder (ASPD; 25.4%) the most common. The overlap between ecstasy use and ASPD is likely due, at least in part, to the common underlying personality trait of sensation seeking, which predicts ecstasy use, conduct disorder and ASPD disorder symptoms. Proportionately more women were diagnosed with depression, generalized anxiety disorder, and posttraumatic stress disorder (PTSD), while proportionately more men were diagnosed with ASPD. Proportionately more nonwhite participants had attention deficit hyperactivity disorder (ADHD). Higher levels of education were associated with proportionately less PTSD, ASPD, and ADHD. Higher frequencies of ecstasy use were associated with proportionately more ASPD and ADHD. Comparing the age of first ecstasy use with the age of onset for selected psychiatric disorders revealed that for most participants, disorders preceded use. Those who had used ecstasy more than 50 times were more likely to have experienced a lifetime disorder.

There is a neurobiological relationship between ecstasy use and depression, as MDMA can cause serotonin depletion as well as serotonergic neurodegradation, which may result in depression among users of the drug. Several studies have examined depressive symptomatology among ecstasy users. They have generally found low levels of depressive symptomatology among study participants. For example, one study found two-thirds of the sample had scores that placed them in the nondepressed/minimal depression category, while 4.7% had scores indicative of severe depression. Men were significantly less likely than women and people who used opioids were significantly more likely than nonusers to have higher levels of depressive symptomatology. Higher lifetime occasions of MDMA use were marginally related to symptoms of serious depression.

It is also possible that those who are susceptible to mental health problems are more vulnerable to using ecstasy. For example, adults who had past year psychiatric symptoms (both depressive and panic symptoms) are twice as likely to be recent-onset ecstasy users compared with those without past year psychiatric symptoms. One suggested mechanism is the desire in those who have mood disorders to regulate mood through the use of mood-altering drugs such as ecstasy (i.e. self-medicate).

Behavioral Correlates

A number of behavioral correlates of ecstasy use have been identified, including engaging in deviant behaviors and having risky sex. In a large epidemiological study, recent-onset ecstasy use was significantly more likely

to occur among adolescents and adults (18–34 years old) who engaged in deviant behaviors (i.e. attacking another person, carrying a handgun, selling drugs, stealing) during the past year compared with those who did not engage in those behaviors. Higher levels of deviancy indicated a higher likelihood of being a recent-onset ecstasy user, and associations were strongest with nonviolent deviant behaviors (e.g. selling illegal drugs and stealing). Associations between deviant behaviors and recent-onset ecstasy use were similar in strength to associations between deviant behaviors and recent-onset cocaine and marijuana use in this study, suggesting the existence of a common pathway from deviance to substance use onset.

Ecstasy use is associated with having more friends who use drugs, particularly among young women. Among college students, ecstasy users were more likely to have multiple sexual partners, spend more time socializing with friends, and spend less times studying compared to those who do not use ecstasy. However, ecstasy users are not necessarily academic underachievers, and their satisfaction with education is not necessarily any different from that of nonusers.

Ecstasy use can be associated with more sexually risky behaviors. Up to one-third of young people who currently use ecstasy have more than one sex partner in a given month, and sexual protection rates can tend to be low in this population. Some of the predictors that have been associated with the likelihood of having multiple sex partners among ecstasy users included being nonwhite, younger age at first sexual experience, using ecstasy for its touch-enhancing qualities, having higher self-esteem, handling disagreements more dysfunctionally, and not being involved in a romantic relationship.

Cognitive/Motivational Correlates

A number of studies have identified reasons why ecstasy users initiate use or maintain using once they have started. Positive attitudes toward drug use are associated with initiation of ecstasy use among adolescents. Common reasons why young people say they started using ecstasy is to experiment, while reason they maintain use tend to be to feel good or high and to have a good time with friends. Other common reasons for using are for their stimulant properties (e.g. to increase energy, to stay awake), to enhance the effect of another drug, to seek insight, or to relax and relieve tension. Less common reasons for using tend to be to get away from their problems, being hooked, to get through the day, or because of anger or frustration.

Motivational factors play a role in maintaining ecstasy use. Psychological models emphasize internal regulatory cues that motivate drug use and play a contributory role in dependence. Motivational cues tend to be strong predictors of both ecstasy use and dependence in users. For example, female ecstasy users report they are more willing to use drugs in the future compared to nonusers. Other cognitive correlates of ecstasy use are low risk perception and high perceived behavioral control of obtaining ecstasy (being able to obtain it in a short period of time).

Environmental/Familial Correlates

A number of parental or familial correlates of ecstasy use have been identified. Ecstasy users report more childhood experiences of physical abuse, emotional neglect, and physical neglect than nonusers of MDMA. Parent drug use has been identified as a significant predictor of child initiation of ecstasy use. Living with both parents and close parental monitoring are negatively associated with ecstasy use initiation, and may be protective against it. Current ecstasy users report experiencing greater difficulties with family relationships than nonusers. Peer correlates of ecstasy use include close associations with deviant peers and drug use by close friends. Current ecstasy users report greater difficulties with peer relationships than nonusers.

Economic Correlates

Some research indicates that monetary and opportunity cost, but not income, significantly predicts ecstasy use.

CONSEQUENCES OF ECSTASY USE

Positive Consequences

MDMA has become a popular drug, in part because of the positive effects that a person may experience within an hour or so after taking a single dose. Those effects include feelings of mental stimulation, emotional warmth, empathy toward others, a general sense of well-being, and decreased anxiety. In addition, users report enhanced sensory perception as a hallmark of the MDMA experience.

Among those who have used ecstasy recently (in the past year), a number of positive consequences have been reported. The most common ones are an increase in energy or stimulant effects, and improved social skills or making new friends, and experiencing everything as more fun. Less common but still reported are weight loss, euphoria, experiencing everything as more intense, bettering a current relationship, decreasing inhibitions, and relaxing or relieving stress.

Physical Consequences

Ecstasy can produce a variety of immediate and longer-term adverse health effects, including nausea, profuse sweating, chills, involuntary teeth clenching, tingling or numbness, and blurred vision (due to pupil dilation). Additional negative symptoms can include memory lapse, shortness of breath, inability to urinate, vomiting, joint stiffness, stomach pains, headaches, heart palpitations, tremors or shakes, weight loss, loss of energy, and muscle aches. Ecstasy overdose can also occur; the symptoms can include high blood pressure, faintness, panic attacks, and in severe cases, loss of consciousness and seizures.

Because of its stimulant properties and the environments in which it is often taken, ecstasy is associated with vigorous physical activity for extended periods. This can lead to one of the most significant, although rare, acute adverse effects: a marked rise in body temperature (hyperthermia) which can in turn result in kidney failure. In addition, dehydration, hypertension, and heart failure may occur. Ecstasy can also reduce the pumping efficiency of the heart, of particular concern during periods of increased physical activity, further complicating these problems. Because MDMA affects monoamine neurotransmitters that play a critical role in sleep and daytime alertness, ecstasy use interferes with sleep. Acute MDMA administration disrupts sleep and rapid eye movement (REM) sleep, specifically, without producing daytime sleepiness in the same was sleep restriction does. Compared with control subjects, recreational ecstasy users show evidence of hyperarousal and impaired REM function.

MDMA is rapidly absorbed into the human bloodstream, but once in the body, MDMA metabolites interfere with the body's ability to metabolize, or break down, the drug. As a result, additional doses of ecstasy can produce unexpectedly high blood levels, which could worsen the cardiovascular and other toxic effects. MDMA also interferes with the metabolism of other drugs, including some of the adulterants that may be found in ecstasy tablets.

Psychological Consequences

Due to the effects MDMA has on monoamines that play a role in mood regulation, ecstasy use has both acute and longer-term effects on psychological health. Some ecstasy users report experiencing visual and auditory hallucinations, depression confusion, anxiety, irritability, paranoia, and loss of sex urge at least one time after ecstasy use. However, the psychological effects of ecstasy use (e.g. depression and anxiety) are thought to be due at least in part to other substance use, which is very common among those who use ecstasy. A limited body of work has found that ecstasy and marijuana users have reported more intense feelings of depression and anxiety than those who use marijuana alone and nondrug users. A growing body of research has shown that although ecstasy users demonstrate elevated levels of psychological symptoms, these symptoms are not necessarily associated with their ecstasy consumption. Instead, other drug use (alcohol, marijuana, opioids, and inhalants) significantly account for these psychological symptoms. A number of studies have pointed specifically to marijuana use to explain elevated depression symptoms and clinical depression among ecstasy users. These studies have shown that those who use ecstasy and marijuana have higher rates of depression than those who use ecstasy use alone, and that ecstasy use, severity of use, or a diagnosis of dependence on ecstasy is not associated with depression without considering the effects of marijuana.

Lifestyle Consequences

Ecstasy use can cause a number of lifestyle consequences. In one study, 40% of recent ecstasy users reported trouble maintaining their usual daily activities, while 20% reported experiencing financial and work trouble as a result of using ecstasy. Problems with aggression were infrequent after club drug use. A small group of participants indicated that they got into verbal arguments sometimes (14%) or half the time (2%) after ecstasy use. A small percentage of individuals indicated that they ever had been involved in physical (14%) or sexual (2%) aggression after ecstasy use. Ecstasy use has been associated with low academic achievement in some epidemiological studies.

Ecstasy use also raises concerns because of its association with risky driving. Research indicates that ecstasy use impairs the perception of distance between cars, even in recently abstinent ecstasy users. Other work shows that while ecstasy use may not impair driving performance, it leads to elevated risk taking in driving situations.

Risky Sex

Ecstasy use can lead to risky sexual behaviors. Among females, findings are mixed; however, ecstasy use has been associated with having unprotected intercourse and engaging in sex acts with partners that they would not have otherwise. Among young adult ecstasy users, nearly one-third have had more than one sex partner during the preceding month, and sexual protection rates have tended to be low in some studies. Predictors of having multiple sex partners include being nonwhite, knowing someone who was HIV-positive, younger age of first sexual experience, using ecstasy

for its touch-enhancing qualities, higher self-esteem, handling disagreements more dysfunctionally, and not being involved in a romantic relationship. In contrast, heavy ecstasy users are more likely to have been tested for HIV than nonheavy users, but are also more likely to perceive no chance of contracting HIV. Although ecstasy use is associated with more HIV risk behaviors, it is not associated with greater likelihood of being diagnosed with a sexually transmitted disease.

A number of studies have examined whether ecstasy use is associated with more HIV risk behaviors among populations that are at high risk for contracting HIV. For example, among MSM, substance use before or during sex has not been associated with risk with HIV-negative partners, but has been associated with risk with partners who are HIV-positive and unknown serostatus. Further, among MSM assessed from 2001 to 2007, a greater number of HIV-positive (versus HIV-negative) men reported recent drug use, including ecstasy, across years. Among Hispanic MSM, having a higher number of sex partners, having higher social isolation scores, and having engaged in unprotected receptive anal intercourse have been associated with club drug use. Club drug users tend to have more sex partners than nonusers of club drugs and among those who do use club drugs, those who use more severely are more likely to have sex under the influence of club drugs compared to those who use less severely. Among Asian/Pacific Islander MSM, ecstasy is the most common drug used in conjunction with sex and an association has been demonstrated between unprotected anal intercourse and being high or buzzed on ecstasy, but not alcohol, marijuana, GHB, or crystal methamphetamine.

Ecstasy may also contribute to sexual side effects in men and women. For both, there may be a decreased ability to achieve orgasm if engaging in sex while on the drug, and for men, inability to maintain an erection. Efforts to achieve orgasm might include prolonged sexual encounters or rougher encounters, both of which can potentiate opportunities for a condom to break or other forms of genital irritation to occur, which increases the risk for sexually transmitted infections and HIV transmission. Although the physical ability to have sex may be diminished while high on ecstasy, sensual/sensory aspects of the drug can create a desire for touch and eroticism (e.g. kissing, fondling, genital stimulation). In one study of heterosexual female ecstasy users, this was described as sensuality as opposed to sexuality.

Neurocognitive Consequences

Numerous studies on animals have demonstrated that MDMA can damage serotonin-containing neurons; some of these studies have shown these effects to be long lasting. This suggests that such damage may occur in humans as well; however, measuring serotonin damage in humans is more difficult. Studies have shown that some heavy ecstasy users experience long-lasting confusion, depression, and selective impairment of working memory and attention processes. Such memory impairments have been associated with a decrease in serotonin metabolites or other markers of serotonin function. Functional magnetic resonance imaging studies on MDMA users have shown changes in brain activity in regions involved in cognition, emotion, and motor function. However, improved imaging technologies and more research are needed to confirm these findings and to elucidate the exact nature of the effects of MDMA on the human brain.

In 2007, a meta-analysis reviewed the literature on the association between MDMA misuse and neurocognition. MDMA use was found to be associated with neurocognitive deficits in attention/concentration, verbal and nonverbal learning and memory, psychomotor speed, and executive systems functioning. As with the psychological symptoms of ecstasy use, researchers have suggested that the neurocognitive deficits demonstrated in ecstasy users can be attributed to other substances that are commonly used in this population such as marijuana and alcohol. Some work has demonstrated that ecstasy and marijuana users are no more susceptible to these neurocognitive deficits than marijuana users alone. For example, one study has shown that ecstasy and other drug users showed comparable patterns of decision-making and impulsivity which was worse in both groups compared to control subjects who did not use ecstasy. Another study demonstrated that memory and executive functioning was similar in those who use ecstasy and those who use ecstasy and cannabis, and that both functions were poorer in these groups compared to controls. This work suggests that some cognitive impairment attributed to ecstasy is more likely due to concurrent marijuana use. However, at least one controlled study has demonstrated that ecstasy and marijuana users perform significantly worse in learning and memory compared to controls and marijuana users and that this worse performance can be attributed to differences in brain activity among ecstasy users compared to those who only use marijuana.

Effects on the Developing Fetus

Given that most ecstasy users are young and in their reproductive years, it is possible that some female users may be pregnant when they take ecstasy, either inadvertently or intentionally. Animal studies have shed light on the potential adverse effects of MDMA on the developing fetus. Behavioral studies in animals have found significant adverse effects on tests of learning and memory from exposure to MDMA during

a developmental period equivalent to the third trimester in humans. However, the effects of MDMA on animals earlier in development are unclear; therefore, more research is needed to determine what the effects of MDMA and ecstasy are on the developing human nervous system.

TREATMENT FOR ECSTASY USE

There are no specific treatments for ecstasy use and associated problems. However, harm reduction efforts have directly targeted ecstasy users by providing pill testing services at nightclubs and dance venues. For example, DanceSafe is a nonprofit organization that supplies volunteers at raves and nightclubs to provide information on ecstasy and other club drugs, including the dangers of using the drug and test pills for chemical makeup/harmful adulterants.

The most effective treatments for drug abuse and addiction in general are cognitive behavioral interventions that are designed to help modify the patient's thinking, expectancies, and behaviors related to their drug use and to increase skills in coping with life stressors. Drug abuse recovery support groups may also be effective in combination with behavioral interventions to support long-term, drug-free recovery. There are currently no pharmacological treatments for addiction to MDMA.

SUMMARY

In this chapter, we reviewed the history and patterns of MDMA/ecstasy use, along with its correlates and consequences. Although it is an illegal Schedule 1 drug, ecstasy remains popular, particularly among young adults and has been connected most strongly with the dance/rave scene (although the drug proliferates within other subcultures as well). In users, the drug produces intense feelings of euphoria, coupled with increased energy and feelings of sensuality and belonging. These factors are likely significant motivators for use in spite of its negative effects, which include acute impairment of memory and cognition, and can lead to confusion and feelings of depression and anxiety. It seems symptoms of dependence are common among users, though physical addiction is typically not associated with the drug. As a result, there are limited options available for treatment, which has perhaps furthered the growth of community-based groups providing harm reduction alternatives for users.

SEE ALSO

Methamphetamine Addiction, Hallucinogens, Ketamine, Epidemiology of Adolescent and Young Adult Alcohol, Tobacco, and Drug Use and Misuse in the United States, Alcohol and Drug Use in Sexual Minority Youth and Young Adults (lesbian, gay, bisexual, transgender)

Glossary

Amphetamine (speed, bennies, black beauties, crosses, hearts, LA turnaround, truck drivers, and uppers) amphetamines are stimulants that often come in pill form and are sometimes prescribed by doctors for medical problems, most commonly attention deficit hyperactivity disorder (ADHD). Amphetamines can also be abused; that is, used in a way other than as prescribed (e.g. crushed and snorted) or used by someone without a prescription.

ADHD attention deficit hyperactivity disorder.

ASPD antisocial personality disorder.

Club drugs substances including ecstasy (MDMA), γ-hydroxybutyrate (GHB), ketamine, Rohypnol, methamphetamine, and lysergic acid diethylamide (LSD), commonly used in all-night dance parties such as raves. Use of these substance proliferated through the 1990s, and they remain widely used in nightclubs as well as other settings (bars, house parties).

Dopamine dopamine is a catecholamine neurotransmitter that is produced in several areas of the brain, including the substantia nigra and the ventral tegmental area. Dopamine is also a neurohormone released by the hypothalamus. Its main function as a hormone is to inhibit the release of prolactin from the anterior lobe of the pituitary. Dopamine affects brain processes that control movement, emotional response, and ability to experience pleasure and pain. Regulation of dopamine plays a crucial role in our mental and physical health. Neurons containing dopamine are clustered in the substantia nigra (midbrain). In Parkinson disease, the dopamine-transmitting neurons in this area die. As a result, the brain of people with Parkinson disease contains almost no dopamine. The drug L-DOPA can be converted in the brain to dopamine, thereby reducing these symptoms.

GHB (Xyrem, Liquid E, Liquid Ecstasy, Liquid G, Gina) γ-Hydroxybutyrate (GHB), one of the club drugs, is a central nervous system depressant that was approved by the US Food and Drug Administration (FDA) in 2002 for use in the treatment of narcolepsy. It is odorless and colorless but has a strong flavor, so it is frequently taken with alcohol or other beverages that mask the taste. It has been used to commit sexual assaults (date rape). In addition to intoxicating effects, GHB has anabolic effects (it stimulates protein synthesis) and has been used by bodybuilders to aid in fat reduction and muscle building.

Ketamine (Special K) ketamine is one of the drugs considered as club drugs. It is a dissociative esthetic, used in veterinary practice, that is usually snorted or injected by humans. It is traditionally in liquid form, and is often cooked down to a solid for nasal ingestion.

IDU intravenous drug user.

LSD lysergic acid diethylamide.

MDMA 3,4-Methylenedioxymethamphetamine.

Methamphetamine (speed, meth, chalk, tina, ice, crystal, crank, glass, fire, go fast) methamphetamine is a stimulant, originally derived from amphetamine, which comes in powder form and can be ingested orally, nasally, or injected intravenously. It is often made in large laboratories called superlabs, but is also sometimes made in small laboratories using inexpensive over-the-counter and often toxic ingredients (such as drain cleaner, battery acid, and antifreeze).

Methylenedioxymethamphetamine (MDMA, ecstasy, E, X, XTC) one of the most common club drugs, MDMA is a synthetic drug that has both stimulant and psychoactive properties. Although commonly associated with all-night dance parties (raves), its use has expanded to other settings, especially among young people. MDMA acts primarily on the serotonergic (5-HT) system, involved in many psychological states including mood and sleep.

Monoamine neurotransmitters monoamine neurotransmitters in biochemistry include organic compounds with only one amino group. The most important members in this group are the neurotransmitter group catecholamines (including dopamine, epinephrine, and norepinephrine) and the indoleamine serotonin. Monoamines seem to contribute to stable moods, and an excess or deficiency of monoamines seems to cause or result from several mood disorders.

MSM men who have sex with men.

Norepinephrine norepinephrine is a catecholamine neurotransmitter and also a stress hormone. As a stress hormone, norepinephrine affects parts of the brain, such as the amygdala, where attention and responses are controlled. Along with epinephrine, norepinephrine also underlies the fight-or-flight response, directly increasing heart rate, triggering the release of glucose from energy stores, and increasing blood flow to skeletal muscle.

NSDUH National Survey on Drug Use and Health.

PTSD posttraumatic stress disorder.

REM rapid eye movement.

Seroconversion seroconversion is the development of detectable specific antibodies to microorganisms in the blood serum as a result of infection or immunization. The word is often used in reference to testing for anti-HIV antibodies, referring to the process of becoming HIV-positive.

Serotonin serotonin is a monoamine neurotransmitter commonly found in the gastrointestinal system (gut) and the central nervous system. In the brain, serotonin neurotransmitters are involved in the regulation of mood, appetite, sleep, and muscle contraction. Serotonin also has some cognitive functions, including in memory and learning. Modulation of serotonin at synapses is thought to be a major action of several classes of pharmacological antidepressants.

Further Reading

Grov, C., Kelly, B.C., Parsons, J.T., 2009. Polydrug use among club-going young adults recruited through time-space sampling. Substance Use and Misuse 44 (6), 848–864.

Kalechstein, A.D., De La Garza 2nd, R., Mahoney 3rd, J.J., Mahoney 3rd, J.J., Fantegrossi, W.E., Newton, T.F., 2007. MDMA use and neurocognition: a meta-analytic review. Psychopharmacology 189 (4), 531–537.

Leung, K.S., Cottler, L.B., 2008. Ecstasy and other club drugs: a review of recent epidemiologic studies. Current Opinion in Psychiatry 21, 234–241.

Maxwell, J.C., 2005. Party drugs: properties, prevalence, patterns, and problems. Substance Use and Misuse 40, 1203–1240.

O'Leary, G., Nargiso, J., Weiss, R.D., 2001. 3,4-Methylenedioxymethamphetamine (MDMA): a review. Current psychiatry Reports 3 (6), 477–483.

Parsons, J.T., Grov, C., Kelly, B.C., 2009. Club drug use and dependence among young adults recruited through time-space sampling. Public Health Reports 124 (2), 246–254.

Ramo, D.E., Grov, C., Delucchi, K., Kelly, B.C., Parsons, J.T., 2010. Typology of club drug use among young adults recruited using time-space sampling. Drug and Alcohol Dependence 107, 119–127.

Romanelli, F., Smith, K.M., Pomeroy, C., 2003. Use of club drugs by HIV-seropositive and HIV-seronegative gay and bisexual men. Topics in HIV Medicine 11 (1), 25–32.

Relevant Websites

http://www.dancesafe.org/ – DanceSafe, drug information and harm reduction resources.

https://dawninfo.samhsa.gov/ – Drug Abuse Warning Network (DAWN) data.

http://www.ecstasydata.org/ – providing access to lab testing results for street ecstasy tablets.

http://monitoringthefuture.org – Monitoring the Future a continuing study of American youth.

http://www.nida.nih.gov/drugpages/mdma.html – National Institute on Drug Abuse (NIDA): Ecstasy (MDMA).

http://oas.samhsa.gov/nsduhLatest.htm – SAMHSA latest National Survey on Drug Use and Health results.

CHAPTER

73

Inhalants

Li-Tzy Wu[*], *George E. Woody*[$]

[*]Duke University Medical Center, Durham, NC, USA,

[$]University of Pennsylvania and Treatment Research Institute, Philadelphia, PA, USA

OUTLINE

INTRODUCTION

In the United States, about 16% of eighth graders, 13% of tenth graders, and 10% of twelfth graders had ever used inhalants in their lifetime. In the general population, about 10% of American adults aged 18 years or older reported a history of inhalant use (13.4% of men; 6.3% of women). Inhalant use confers a high risk for significant morbidity and mortality. It is an important public health concern worldwide because household products and office supplies subject to inhalant use or misuse are omnipresent, easily accessible (from homes, offices, and a variety of stores), relatively inexpensive, conveniently packaged, and quickly produce the onset of high and toxic effects. Inhalants are particularly attractive to children, youth, and socioeconomically disadvantaged populations. This chapter summarizes current knowledge about inhalant use, including patterns of use; incidence and prevalence of use; assessment and diagnosis for inhalant use disorders according to the *Diagnostic and Statistical Manual of Mental Disorders Fourth Edition* (DSM-IV); comorbid psychiatric disorders, medical consequences, and mortality associated with inhalant use; and prevention and treatment. This review focuses on research data from major studies that include a large sample size (to ensure the generalizability of study findings) and have adequate coverage of respondents from diverse racial/ethnic groups and geographic regions.

WHAT IS INHALANT USE?

Inhalant use is the deliberate ingestion of volatile substances via (1) sniffing or snorting fumes from containers; (2) spraying aerosols directly into the nose

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00073-5

or mouth; (3) sniffing or inhaling fumes from substances sprayed or deposited inside a plastic or paper bag (i.e. bagging); (4) huffing from an inhalant-soaked rag stuffed in the mouth; or (5) inhaling from balloons filled with nitrous oxide to induce a psychoactive or mind-altering effect. Based on the profiles of pharmacological and behavioral effects, three main groups of inhalants are subject to these use patterns: (1) volatile solvents, fuels, and anesthetics; (2) nitrous oxide; and (3) volatile alkyl nitrites (e.g. amyl, butyl, and isobutyl nitrites). Some commonly used sources of inhalants include adhesives (e.g. airplane glue, rubber cement), aerosols (e.g. spray paint, hair spray, deodorant, room freshener, analgesic spray, asthma spray), cleaning agents (e.g. dry cleaning chemicals, spot remover, degreaser), solvents (e.g. nail polish remover, paint thinner, correction fluid, fuel gas, lighter, fire extinguisher), and gasoline. Any of these substances can produce a feeling of euphoria or high.

PATTERNS OF INHALANT USE

Surveys of the General Population

According to results from the National Survey on Drug Use and Health (NSDUH), a national survey of a representative sample of noninstitutionalized civilian residents aged 12 years and older in the United States (Table 73.1), the most commonly used inhalants are glue, shoe polish, toluene, gasoline, lighter fluid, spray paints, correction fluid, degreaser, nitrous oxide, and whippets for adolescents under age 18 years; nitrous oxide and whippets for young adults aged 18–25 years; and amyl nitrite (poppers or rush) for adults aged 26 years or older.

TABLE 73.1 Types of Inhalants Used Among Adolescents and Adults from the Most Commonly Used Inhalants to the Least Commonly Used Inhalants

Types of inhalants used among adolescents aged 12–17 years	Types of inhalants used among adults aged 18 years and older
Glue, shoe polish, toluene	Nitrous oxide, whippets
Gasoline, lighter fluid	Amyl nitrite, poppers, rush
Spray paints	Glue, shoe polish, toluene
Correction fluid, degreaser	Gasoline, lighter fluid
Nitrous oxide, whippets	Correction fluid, degreaser
Aerosol sprays	Paint solvents, lacquer thinner
Amyl nitrite, poppers, rush	Aerosol sprays
Paint solvents, lacquer thinner	Spray paints
Lighter gases, butane, propane	Ether, halothane, anesthetic
Ether, halothane, anesthetic	Lighter gases, butane, propane

Notes: These data are based on research findings from Wu et al. (2004) and Wu and Ringwalt (2006).

Results from the NSDUH show important age and gender differences in types of inhalants used (Table 73.2). Boys are more likely than girls to use gasoline, lighter fluid, nitrous oxide, and whippets, while girls are more likely than boys to use glue, shoe polish, toluene, spray paints, correction fluid, degreaser, and aerosol sprays. Among adults, men are more likely than women to use glue, shoe polish, toluene, gasoline, lighter fluid, paint solvents, lacquer thinner, lighter gases, butane, and propane.

Compared with adults, adolescents tend to use more types of inhalants and have a more frequent pattern of use. Approximately 50% of lifetime adolescent inhalant users in the United States have used multiple inhalants, while about 38% of lifetime adult inhalant users have a history of using multiple inhalants. In the sample of recent (past year) inhalant users, one-fifth of adolescents are weekly users of inhalants compared with about one-tenth of adults.

The National Poison Data System

Data from poison control centers provide additional information about individuals who seek help or treatment for inhalant use and the demographic characteristics of subgroups affected by use of harmful inhalants. According to data from the National Poison Data System (NPDS), the vast majority of inhalant cases are adolescents aged 12–17 years. Propellants (e.g. computer and electronics duster sprays, fluorocarbons), gasoline, and paint are the products most frequently implicated in inhalant cases reported to the NPDS. While there have been fewer changes in use of gasoline and paint products during the past few years, there has been a significant and steady increase in the use of propellants (computer and electronics duster sprays) over time. The NPDS data show that gasoline is the product most commonly used by children and adolescents under the age of 14 years, while propellants are the category of use most frequently cited by adolescents, especially older adolescents.

INCIDENCE AND PREVALENCE OF INHALANT USE

Incidence

Research findings on onset or incidence (i.e. new cases) of inhalant use can contribute important clues about the factors associated with initial use and can be

TABLE 73.2 Gender Differences in Types of Inhalants Used

Adolescents aged 12–17 years		Adults aged 18 years or older	
Glue, shoe polish, toluene	Girls > boys	Glue, shoe polish, toluene	Men > women
Spray paint	Girls > boys	Gasoline, lighter fluid	Men > women
Correction fluid, degreaser	Girls > boys	Paint solvent, lacquer thinner	Men > women
Aerosol spray	Girls > boys	Lighter gas, butane, propane	Men > women
Gasoline, lighter fluid	Boys > girls	–	–
Nitrous oxide, whippets	Boys > girls	–	–

Notes: These estimates are based on the results from Wu et al. (2004) and Wu and Ringwalt (2006).

useful for planning the timing and foci of prevention interventions. Onset of use can occur in children as young as 5–6 years of age. Research reports from the Monitoring the Future (MTF) study, a series of classroom surveys of eighth, tenth, and twelfth graders in the United States, show that inhalant use occurs early, with peak initiation rates in grades six to nine, and decreases with increasing age. However, it is important to note that school dropouts are not covered by a school-based survey like the MTF and that school dropouts in general have a high rate of illicit or nonmedical drug use, including inhalant use.

Surveys of community-based or noninstitutionalized populations that include school dropouts, such as the NSDUH, reveal that the risk of initiating inhalant use is not limited to early adolescence. Among adolescent inhalant users, the vast majority (80%) initiate inhalant use before the age of 15 years, and the remaining adolescents start inhalant use between 15 and 17 years of age. In the United States, national sample of adult inhalant users, about one-third initiate use between age 15 and 17 years, and 44% start, use in adulthood. These onset patterns are in line with age-related differences in types of inhalants used. Specifically, adolescents are more likely than adults to use readily accessible household products such as glue, shoe polish, gasoline, or spray paints, as inhalants of choice. By comparison, adults tend to initiate inhalant use within the context of a severe pattern of polysubstance use and sexual risk taking, and their inhalants of choice are likely to include nitrous oxide (whippets) and amyl nitrite (poppers or rush for sex-related activities), which are much more commonly used by illicit drug users and homosexual or bisexual individuals than by the general population.

Estimates from the 2007–2008 NSDUH show that approximately 730 000–770 000 Americans initiate inhalant use (first-time use) during a 12-month period, which represents slightly more than one-third of individuals who used inhalants in the past year. The level of perceived risk of using a psychoactive drug once or twice is generally considered an indicator for drug use (i.e. greater perceived risk by adolescents results in less psychoactive drug use). In the MTF, eighth and tenth graders are asked questions about the degree of risk they associate with using an inhalant once or twice. Unfortunately, there has been a decline in the perceived risk associated with inhalant use since 2000, and a relatively low proportion of eighth and tenth graders (less than 40% in 2008) reported that there is a "great risk" in using an inhalant once or twice. The hazards of inhalant use were communicated during the mid-1990s via an anti-inhalant advertising initiative launched by the Partnership for a Drug-Free America. The decline in perceived risk for using inhalants may be related to a generational forgetting of the hazards of inhalant use as younger cohorts were not exposed to these messages. This steady decline in perceived risk is worrisome and requires close monitoring of national trends and associated problems because inhalant use in adolescence is a robust predictor for serious substance abuse problems, including polysubstance use, heroin use, drug abuse or dependence, and injection drug use.

Prevalence

Of the nine main categories of illicit or nonmedical psychoactive drug use assessed by the MTF, inhalants are the second most widely used class of drugs among eighth and tenth graders (after marijuana) and are the third most widely used among twelfth graders (after marijuana and prescription opioids) with a lifetime rate similar to amphetamine use. In 2008, about 16% of eighth graders, 13% of tenth graders, and 10% of twelfth graders had ever used inhalants in their lifetime (Table 73.3).

School-based MTF surveys have generally found a higher prevalence of lifetime inhalant use among eighth graders than among older students, a trend that stands in contrast with lifetime prevalence rates of other categories of illicit or nonmedical drug use, which usually increase with age. This unique finding may be related to an association between early inhalant use and dropping out of school; that is, adolescents who

TABLE 73.3 Lifetime Prevalence (%) of Inhalant Use in Eighth, Tenth, and Twelfth Grades: Results from the MTF

Grade	2000	2001	2002	2003	2004	2005	2006	2007	2008
8th	17.9	17.1	15.2	15.8	17.3	17.1	16.1	15.6	15.7
10th	16.6	15.2	13.5	12.7	12.4	13.1	13.3	13.6	12.8
12th	14.2	13.0	11.7	11.2	10.9	11.4	11.1	10.5	9.9

Notes: These data are based on research findings from Johnston, L. D., O'Malley, P. M., Bachman, J. G., and Schulenberg, J. E. (2009). *Monitoring the Future national results on adolescent drug use: overview of key findings, 2008*. Bethesda, MD: National Institute on Drug Abuse.

use inhalants tend to drop out of school and are thus not included in school-based MTF surveys. Alternatively, older adolescents may be differentially more likely to under report inhalant use than use of other psychoactive drugs due to the perception that inhalants are kids' drugs. The results of inhalant use for 10th and 12th graders are considered underestimated. This is an area warranting further research.

Adolescents who are native American, native Alaskan, or of multiple races have a particularly high rate of inhalant use compared with adolescents of other racial or ethnic backgrounds; adolescents who are Asian and black have a very low rate of inhalant use. Among native American youths, inhalants are often the first psychoactive drug used. Earlier research suggests that boys are more likely than girls to use inhalants; more recent research data, however, indicate a more similar rate of inhalant use among genders. Because many adolescent girls of childbearing age are within the population of young inhalant users and there are possible adverse effects of maternal inhalant use on fetuses, the pattern and extent of inhalant use in female adolescents and youths warrant close scrutiny. Adolescent inhalant use also increases the risk for conduct problems, binge drinking, use of multiple drugs, heroin use, injection drug use, and depression.

In the general population, fewer adults than adolescents use inhalants. According to data from the NSDUH, approximately 10% of noninstitutionalized civilian adults in the United States have used inhalants in their lifetime (13.4% in men; 6.3% in women), and about 0.5% (5% of lifetime inhalant users) used inhalants in the past year. Women and blacks have a lower rate of inhalant use than men and other racial or ethnic groups. Additional characteristics associated with inhalant use include younger age (less than age 34 years), being single, receipt of mental health treatment in the past year, self-reported serious psychological distress, involvement with the criminal justice system, alcohol abuse or dependence, and past year drug use (especially polydrug use).

Nitrite Inhalant Use

Of all inhalants used, use of nitrite inhalants (amyl, butyl, and isobutyl nitrites or poppers) is associated with an elevated risk for infections and transmission of sexually transmitted diseases (e.g. human immunodeficiency virus (HIV)) because users often are polysubstance users or engage in unprotected or risky sexual behaviors. Inhaling nitrites can dilate blood vessels, increase the heart rate, and produce a sensation of heat and excitement that may last for several minutes. Since the early 1960s, nitrite inhalants have been used by individuals to enhance sexual activities or to get high. Although the rate of nitrite inhalant use is low in the general population, nitrite inhalant use is relatively common among drug abusers (especially individuals in addiction treatment) and homosexual or bisexual individuals, and its use is associated with risky sexual behaviors, illicit drug use, transmission of sexually transmitted diseases, suicide attempts, and drug-related overdose.

Nitrite inhalants are the primary inhalant used by adults, and use in adolescents is less frequent. In the MTF, approximately 1% of twelfth graders between 2007 and 2008 used nitrite inhalants in their lifetime. In the NSDUH, about 1.5% of noninstitutionalized adolescents aged 12–17 years (including school dropouts) have ever used nitrite inhalants in their lifetime. While use in the general population of adolescents is infrequent, the rate of lifetime nitrite inhalant use increases to 15% among adolescents who have an alcohol or drug dependence or who used multiple drugs in the past year. Nitrite inhalant use in adolescents can be considered a marker for serious psychiatric problems in that users not only tend to engage in delinquent activities, use multiple types of inhalants and other substances, and have alcohol/drug dependence, but also exhibit mental health problems and have received mental health treatment in the past year. Young nitrite inhalant users thus represent a subgroup of highly troubled youth in need of help. Additional characteristics associated with nitrite inhalant use in adolescents include older age (15–17 years), white race, more than one race, and residence in a nonmetropolitan area.

ASSESSMENT AND DIAGNOSIS FOR INHALANT USE DISORDERS

Assessments and diagnosis for DSM-IV inhalant abuse and dependence can be difficult since an accurate drug use history is not always possible to obtain and routine urine drug screening will not detect inhalants. Potential signs of inhalant use can be subtle and are summarized in Table 73.4.

TABLE 73.4 Potential Signs of Inhalant Use

Odors or chemical smells on the breath or clothes
Residue of the substance on clothing or skin (paint or glitter on face or hands)
Burns
Cans of gasoline or spray paint under an adolescent's bed
Rashes around nose or mouth (e.g. glue sniffer's rash)
Red or irritated eyes, throat, and nose
Trauma or injuries
Nonspecific respiratory problems (e.g. coughing, sinus discharge, dyspnea, or rale)
Headache or general weakness
Abdominal pain, nausea, or vomiting

problems), and the user does not meet the criteria for inhalant dependence in the past year (Table 73.6). DSM-IV inhalant dependence includes inhalant users who show a maladaptive pattern of inhalant use leading to clinically significant impairment or distress, as manifested by having three or more of the six dependence criteria (i.e. tolerance, taking larger amounts over a longer period of time, inability to cut down, a great deal of time spent in using or recovering from its effects, important activities given up, and continued use despite resulting medical or psychological problems) that occur during a 12-month period.

It is important to note that inhalant withdrawal symptoms (e.g. sleep disturbances, tremor, irritability, diaphoresis, nausea, and fleeting illusions) may occur within 1–2 days after cessation of use and may last for 2–5 days. However, because of a lack of research data

According to the DSM-IV (Table 73.5), inhalant intoxication is defined as the presence of clinically maladaptive behavioral or psychological changes that develop during or shortly after intentional use of or exposure to volatile inhalants (e.g. dizziness, nystagmus, lack of coordination, slurred speech, unsteady gait, lethargy, depressed reflexes, psychomotor retardation, tremor generalized muscle weakness, blurred vision or diplopia, stupor or coma, and euphoria) and that are not due to a general medical condition and are not better accounted for by another mental disorder.

Inhalant abuse as defined by the DSM-IV as a maladaptive pattern of inhalant use leading to clinically significant impairment or distress, as manifested by having one or more of the four abuse criteria (i.e. role interference, hazardous use, legal problems, and relation

TABLE 73.5 DSM-IV Criteria for Inhalant Intoxication

A	Recent intentional use or short-term high-dose exposure to volatile inhalants (excluding anesthetic gases and short-acting vasodilators)
B	Clinically maladaptive behavioral or psychological changes that developed during or shortly after use of or exposure to volatile inhalants
C	Two or more of the following signs, developing during or shortly after inhalant use or exposure: dizziness, nystagmus, incoordination, slurred speech, unsteady gait, lethargy, depressed reflexes, psychomotor retardation, tremor, generalized muscle weakness, blurred vision or diplopia, stupor or coma, euphoria
D	The symptoms are not due to a general medical condition and are not better accounted for by another mental disorder

Note: These criteria are based on the DSM-IV, Text Revision (American Psychiatric Association, 2000).

TABLE 73.6 DSM-IV Criteria for Inhalant Use Disorders

Inhalant abuse	Inhalant dependence
A. A maladaptive pattern of inhalant use leading to clinically significant impairment or distress, as manifested by one or more of the following, occurring within a 12-month period: **(1)** Recurrent inhalant use resulting in a failure to fulfill major obligations at work, school, or home **(2)** Recurrent inhalant use in situations in which it is physically hazardous **(3)** Recurrent inhalant-related legal problems **(4)** Continued inhalant use despite having persistent or recurrent social or interpersonal problems caused or exacerbated by the effects of the inhalant **B.** The symptoms have never met the criteria for inhalant dependence for this class of substance	**A.** A maladaptive pattern of inhalant use leading to clinically significant impairment or distress, as manifested by three or more of the following, occurring at any time in the same 12-month period: **(1)** Tolerance as defined by the need for markedly increased amounts of the inhalant to achieve intoxication or desired effects, or markedly diminished effects with continued use of the same amount of the inhalant **(2)** An inhalant is often taken in larger amounts or over longer periods than was intended **(3)** There is a persistent desire or unsuccessful effort to reduce or control inhalant use **(4)** A great deal of time is spent in activities necessary to obtain the inhalant, use it, or recover from its effects **(5)** Important social, occupational, or recreational activities are given up or reduced because of inhalant use **(6)** Inhalant use is continued despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by it

Note: These criteria are based on the DSM-IV, Text Revision (American Psychiatric Association, 2000).

to establish an inhalant withdrawal syndrome, the withdrawal criterion is not included for DSM-IV inhalant dependence. This may change in the future as more empirical data for an inhalant withdrawal syndrome accumulate.

PREVALENCE AND CORRELATES OF INHALANT USE DISORDERS

National surveys of the US general population show that 10.6% of adolescents aged 12–17 years who used inhalants in the past year reported a pattern of inhalant use problems consistent with the DSM-IV criteria for a current inhalant use disorder and that 7.7% of adults aged 18 years or older who used inhalant in the past year reported a pattern of inhalant use problems consistent with the DSM-IV criteria for a current inhalant use disorder.

Inhalant abuse and dependence, as well as cessation from inhalant use, have received comparatively less research attention than inhalant use. Inhalant use was not included in the Diagnostic Interview Schedule of the National Institute of Mental Health-Epidemiologic Catchment Area (NIMH ECA) study. One publication from the 1990 National Comorbidity Survey (NCS) describes the prevalence of lifetime inhalant dependence (0.3%) in the general population. More recent research data suggest that use of multiple inhalants in adolescents is common (about 50% of users) and that many have used before the age of 14 years, findings that point toward the need to identify those at risk for progressing from use to abuse or dependence. Research data from the NSDUH show that 0.2% of adolescents aged 12 years or older met the DSM-IV criteria for current inhalant abuse, and another 0.2% met the criteria for inhalant dependence in the past year. In the subsample of adolescents who used inhalants in the past year (Table 73.7), 6% met the DSM-IV criteria for current inhalant abuse and an additional 4% for inhalant dependence. Therefore, about one in ten adolescent inhalant users have an inhalant use disorder. It is also important to note that these estimates of inhalant use disorders are all based on self-reported data from survey respondents.

Different characteristics are associated with inhalant abuse and dependence in adolescents. Characteristics associated with increased odds of inhalant abuse include engaging in multiple delinquent activities, history of incarceration, and use of multiple inhalants. Characteristics associated with increased odds of inhalant dependence include early onset of first inhalant use, using inhalants weekly, history of foster care placement, receipt of mental health treatment in the past year, and the presence of other drug abuse or dependence.

Data from the NSDUH provide a unique opportunity to compare inhalant use disorders in adults against estimates in adolescents. In the national sample of adults aged 18 years or older, 0.04% met the DSM-IV criteria for current inhalant abuse (0.03%) or dependence (0.01%) in the past year. In the subsample of adults who used inhalants in the past year, 8% met the DSM-IV criteria for current inhalant abuse (7%) or dependence (1%). Although adult men have a higher rate of inhalant use than adult women, there are no gender differences in inhalant use disorders. The following groups of adult inhalant users have an elevated rate of inhalant use disorders: adults aged 35–49 years, those who did not complete high school, users of mental health treatment, and weekly inhalant users. Rates of current inhalant use disorders among adolescents and adults are summarized in Table 73.7.

Research data from the 2001–2002 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) – presently the largest study of psychiatric comorbidity among adults in the United States – provide information about lifetime inhalant use disorders among inhalant users. In the NESARC, of respondents aged 18 years or older, 2% were identified as lifetime inhalant users, and 19% of lifetime users met the DSM-IV criteria for an inhalant use disorder in their lifetime. Most adults (88%) with a lifetime inhalant use disorder met the criteria for inhalant abuse but not inhalant dependence, and 12% met the criteria for inhalant dependence irrespective of whether they had

TABLE 73.7 Prevalence of Past Year or Current DSM-IV Inhalant Use Disorders in the General Population

DSM-IV inhalant use disorders (%)	All respondents		Individuals who used inhalants in the past year	
	Aged 12–17 years	Aged 18 years or older	Aged 12–17 years	Aged 18 years or older
Abuse	0.2	0.03	6.3	6.6
Dependence	0.2	0.01	4.3	1.1
Abuse or dependence	0.4	0.04	10.6	7.7

Note: These estimates are based on results from Wu et al. (2004) and Wu and Ringwalt (2006).

ever met criteria for inhalant abuse. The vast majority (94%) of adults with a history of inhalant use had not used inhalants in the past year, and very few (1%) met the DSM-IV criteria for an inhalant use disorder in the past year.

PSYCHIATRIC COMORBIDITIES, MEDICAL CONSEQUENCES, AND MORTALITY

Substance Use and Other Psychiatric Comorbidities

Adolescent inhalant users often have comorbid DSM-IV substance use disorders and other mental health problems, and adolescents who used both inhalants and marijuana represent a severe set of drug users. Compared with adolescents who use inhalants but no marijuana, or adolescents who use illicit or nonmedical drugs other than inhalants and marijuana, adolescents who use both inhalants and marijuana have a particularly high rate of DSM-IV substance use disorders, including alcohol, marijuana, cocaine, hallucinogen, opioid, sedative, stimulant, tranquilizer, and heroin use disorders. Adolescents who use inhalants but no marijuana have rates of alcohol or drug use disorders similar to adolescents who use illicit or nonmedical drugs other than inhalants.

Adolescent inhalant users in the NSDUH are more likely than noninhalant users to have major depression and to participate in treatment for mental or psychological problems. However, empirical data concerning other DSM-IV mental disorders in the general population of adolescents are lacking. Research data from adolescents in treatment for mental or behavioral problems show that those with a history of inhalant use, abuse, or dependence are more likely than those with other problems to have other substance use disorders, major depression, suicide attempts, and a history of physical/sexual abuse and neglect. These findings suggest that adolescent psychiatric patients with a history of inhalant use should be screened carefully for the presence of serious substance abuse or mental health problems.

Similarly, psychiatric disorders among adults with a history of inhalant use also are prevalent. Research data from the 2001–2002 NESARC have documented that approximately 96% of adults with a history of inhalant use met the DSM-IV criteria for a substance use disorder in their lifetime, and that about two-thirds met the DSM-IV criteria for a substance use disorder in the past year. As shown in Table 73.8, among adults with a history of inhalant use, alcohol, marijuana, nicotine, cocaine, hallucinogen, and stimulant use disorders were all more prevalent than inhalant use disorders (19%). Less education, residence in nonmetropolitan areas, early onset of inhalant use, and a history of substance abuse treatment increased the odds of having an inhalant use disorder among adult inhalant users.

TABLE 73.8 Prevalence of Lifetime DSM-IV Psychiatric Disorders Among Adults who Reported a History of Inhalant Use in the General Population

Lifetime DSM-IV disorder	%	Lifetime DSM-IV disorder	%
Nicotine use disorder	58	Any mood disorder	48
Alcohol use disorder	87	Major depression	41
Inhalant use disorder	19	Dysthymia	18
Marijuana use disorder	68	Mania	15
Hallucinogen use disorder	31	Hypomania	8
Stimulant use disorder	28	Any anxiety disorder	36
Sedative use disorder	17	Panic disorder with agoraphobia	4
Tranquilizer use disorder	18	Panic disorder without agoraphobia	14
Opioid use disorder	20	Social phobia	12
Cocaine use disorder	35	Specific phobia	18
Heroin use disorder	5	Generalized anxiety disorder	11
		Antisocial personality	32

Note: These estimates are based on the results from Wu et al. (2008) and Wu and Howard (2007).

Other (nonaddictive) mental disorders, including mood (48%), anxiety (36%), and personality (45%) disorders, also are prevalent among adults with a history of inhalant use. Particularly common lifetime disorders are major depression and antisocial personality disorder. In addition, psychiatric disorders are highly comorbid, with approximately one in seven lifetime inhalant users meeting the DSM-IV criteria for six or more lifetime nonaddictive mental disorders, and more than one in five inhalant users meeting the criteria for three to five of such disorders. Female inhalant users were more likely than male users to meet the criteria for dysthymia and anxiety disorders, while male inhalant users were more likely than female users to have an antisocial personality disorder. Early onset of inhalant use was strongly associated with an increased likelihood of having multiple mental disorders, particularly mood and personality disorders. On average, inhalant users tend to have an earlier age of onset of mood or anxiety disorders than noninhalant users. Regarding the temporal ordering of various disorders, onset of phobia typically precedes the onset

of inhalant use, whereas mood and other anxiety disorders on average are likely to develop subsequent to inhalant use initiation.

Medical Comorbidities and Mortality

Inhalant use is associated with increased morbidity and mortality. Most inhalants are very toxic to organs, and inhalant users are at risk for an array of long-lasting adverse or even fatal medical consequences, including substantial cardiac, renal, hepatic, and neurological morbidity or mortality. The liver and the heart are the organs most commonly affected by volatile inhalants. Sniffing highly concentrated chemicals in solvents or aerosol sprays can induce irregular and rapid heart rhythms, which can result in heart failure and death within minutes of a session of prolonged sniffing (i.e. sudden sniffing death). Inhalation of nitrites can not only increase the risk for HIV-related risky sexual behaviors but also has adverse effects on the immune system by impairing immune functioning and suppressing resistance to infection. In illicit drug users, daily use of nitrite inhalants (poppers) is also associated with overdose mortality.

Inhalant use is particularly devastating to adolescents with respiratory problems because it can not only produce adverse effects on the respiratory system through irritation or inflammation of breathing passages but can also exacerbate existing respiratory conditions. Recent research data from the NSDUH show that 4–5% of adolescents who used inhalants in the past year also had one or more respiratory conditions (pneumonia, bronchitis, asthma, and sinusitis) in the same period of time. These research data, however, cannot discern whether inhalant use leads to respiratory conditions.

Causes of death related to inhalant use may include suffocation, aspiration, choking, accidental injuries (e.g. car accidents, drowning, fire, trauma), or adverse drug–drug interactions; death can result from either acute (e.g. sudden sniffing death syndrome) or delayed (e.g. cardiomyopathy, central nervous system toxicity, hepatocellular carcinoma, renal toxicity) adverse effects. Recent research data from 60 poison control centers (the NPDS) in the United States provide important information about mortality associated with inhalant use. The mortality rate is defined as the number of deaths per 1000 single substance use cases involving the product. Butane (58 deaths per 1000 cases), propane (26 deaths per 1000 cases), and air fresheners (22 deaths per 1000 cases) have the highest mortality rate, followed by nitrous oxide (14 deaths per 1000 cases), carburetor cleaners (9 deaths per 1000 cases), and fluorocarbons/Freon® (9 deaths per 1000 cases).

The vast majority of inhalant cases reported to the NPDS are adolescent boys, suggesting that boys might tend to use more harmful inhalant products or to engage in heavier or more chronic inhalant use. Girls are most likely to use air freshener, hair spray, and nail products (e.g. nail polish and remover), and the latter two categories have a very low rate of fatality from the analysis of the poison control data. As a group, inhalants have a fatality rate of 5.5 deaths per 1000 cases, which is much higher than the fatality rates (less than 1 death per 1000 cases) for cases of pharmaceutical substances or nonpharmaceutical exposures from the NPDS. Thus, inhalants appear to be more lethal to poison control center cases than other substances.

PREVENTION AND TREATMENT

Prevention

Given the serious, or even fatal, medical consequences resulting from inhalant use, and data showing that use can lead to harmful consequences, prevention efforts should begin with a renewed public information campaign as was done in the 1990s. Because household products and office supplies subject to inhalant use and abuse are omnipresent, a broad and concerted approach that incorporates prevention efforts by schools, communities, parents, and health care professionals are recommended to enhance the general population's knowledge of the serious health consequences of inhalant use and to increase early identification of at-risk subgroups or new onset inhalant users to reduce adverse health effects.

Education by means of a school-based drug abuse curriculum is considered an important component for primary prevention of substance use, and efforts should be made to make certain that inhalant use is included. School staff have the opportunity to identify students who show behavioral or psychological risks for inhalant-related problems and to work with parents of at-risk students. Informing parents and at-risk children and youth about the dangers of inhalant use (e.g. sudden death, burns, and serious brain or liver damage) can decrease experimentation with inhalants. Pediatricians or family physicians can also play an important role in prevention efforts (e.g. education about the dangers from inhalant use, early identification of inhalant users, brief intervention, and referrals to appropriate health care providers) because they see children and youth routinely for physical checkups and often have well-established relationships with family members.

Treatment

According to data from 60 poison control centers in the United States, the majority (56%) of inhalant cases were treated in a health care facility (mainly emergency departments), 10% of cases were admitted to critical care units, 7% to noncritical care units, 8% to psychiatric units, and 20% received an unknown level of care. Inhalant abusers often do not seek treatment for inhalant abuse, and only when inhalant use results in life-threatening or serious consequences does the user present to a health care facility. Medical management starts with providing life support to stabilize the patient and address any acute injury or toxicity. Subsequent treatment needs and plans will depend on the history of physical, mental, and substance abuse status. There are presently no effective reversal agents for inhalant intoxication.

Further, little is presently known about substance abuse treatment needs and successful treatment modalities for inhalant users; clinicians rely on available approaches to substance use problems, such as motivational enhancement techniques, cognitive behavioral therapy, family therapy, or 12-step facilitation. As shown from this review, inhalant abusers are not only affected by multiple substance abuse and mental health problems but might also have developed negative physiological and neurological damage from repeated inhalant use. These multiple comorbid problems pose challenges to effective treatment. Unfortunately, treatment for inhalant abuse or dependence is among the least-studied areas of treatment research. There are few programs designed specifically for inhalant abuse treatment. Thus, access to effective care is limited.

In conclusion, inhalant abuse can lead to serious morbidity and mortality. There is a clear need to increase research efforts on effective prevention and treatment approaches specific to addressing inhalant abuse and dependence.

Glossary

DSM-IV *Diagnostic and Statistical Manual of Mental Disorders Fourth Edition.*
MTF monitoring the future study.
NPDS National Poison Data System.
NSDUH National Survey on Drug Use and Health.
NESARC National Epidemiologic Survey on Alcohol and Related Conditions.

Further Reading

American Psychiatric Association, 2000. Diagnostic and Statistical Manual of Mental Disorders, fourth ed. text revision). American Psychiatric Publishing, Washington, DC.

Anderson, C.E., Loomis, G.A., 2003. Recognition and prevention of inhalant abuse. American Family Physician 68, 869–874.

Balster, R.L., 1998. Neural basis of inhalant abuse. Drug and Alcohol Dependence 51, 207–214.

Brouette, T., Anton, R., 2001. Clinical review of inhalants. American Journal on Addictions 10, 79–94.

Center for Substance Abuse Treatment, 2003. Inhalants. Substance Abuse Treatment Advisory 3, 1–7.

Johnston, L.D., O'Malley, P.M., Bachman, J.G., Schulenberg, J.E., 2009. Monitoring the Future national Results on Adolescent Drug Use: Overview of Key Findings, 2008 (NIH Publication No. 09–7401). National Institute on Drug Abuse, Bethesda, MD.

Marsolek, M.R., White, N.C., Litovitz, T.L., 2010. Inhalant abuse: monitoring trends by using poison control data, 1993–2008. Pediatrics 125, 906–913.

Sakai, J.T., Hall, S.K., Mikulich-Gilbertson, S.K., Crowley, T.J., 2004. Inhalant use, abuse, and dependence among adolescent patients: commonly comorbid problems. Journal of the American Academy of Child and Adolescent Psychiatry 43, 1080–1088.

Sharp, C.W., Rosenberg, N.L., 1997. Inhalants. In: Lowinson, J.H., Ruiz, P., Millman, R.B., Langrod, J.G. (Eds.), Substance Abuse: A Comprehensive Textbook, third ed.). Williams & Wilkins, Baltimore, MD, pp. 246–264.

Substance Abuse and Mental Health Services Administration, 2008. The NSDUH Report: Inhalant Use and Major Depressive Episode Among Youths Aged 12–17: 2004 to 2006. Substance Abuse and Mental Health Services Administration, Office of Applied Studies, Rockville, MD.

Substance Abuse and Mental Health Services Administration, 2010. The NSDUH Report: Adolescent Inhalant Use and Selected Respiratory Conditions. Substance Abuse and Mental Health Services Administration, Office of Applied Studies, Rockville, MD.

Williams, J.F., Storck, M., American Academy of Pediatrics Committee on Substance Abuse, American Academy of Pediatrics Committee on Native American Child Health, 2007. Inhalant abuse. Pediatrics 119, 1009–1017.

Woody, G.E., Donnell, D., Seage, G.R., et al., 1999. Non-injection substance use correlates with risky sex among men having sex with men: data from HIVNET. Drug and Alcohol Dependence 53, 197–205.

Wu, L.T., Howard, M.O., 2007a. Psychiatric disorders in inhalant users: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Drug and Alcohol Dependence 88, 146–155.

Wu, L.T., Howard, M.O., 2007b. Is inhalant use a risk factor for heroin and injection drug use among adolescents in the United States? Addictive Behaviors 32, 265–281.

Wu, L.T., Ringwalt, C.L., 2006. Inhalant use and disorders among adults in the United States. Drug and Alcohol Dependence 85, 1–11.

Wu, L.T., Pilowsky, D.J., Schlenger, W.E., 2004. Inhalant abuse and dependence among adolescents in the United States. Journal of the American Academy of Child and Adolescent Psychiatry 43, 1206–1214.

Wu, L.T., Pilowsky, D.J., Schlenger, W.E., 2005. High prevalence of substance use disorders among adolescents who use marijuana and inhalants. Drug and Alcohol Dependence 78, 23–32.

Wu, L.T., Schlenger, W.E., Ringwalt, C.L., 2005. Use of nitrite inhalants ("poppers") among American youth. Journal of Adolescent Health 37, 52–60.

Wu, L.T., Howard, M.O., Pilowsky, D.J., 2008. Substance use disorders among inhalant users: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Addictive Behaviors 33, 968–973.

Relevant Websites

http://www.inhalant.org/aboutus/ – The Alliance for Consumer Education: Inhalant Abuse Prevention Program.
http://www.aapcc.org/dnn/default.aspx – The American Association of Poison Control Centers (AAPCC).

http://www.drugfreeworld.org/about-us/about-the-foundation.html – The Foundation for a Drug-Free World.

http://monitoringthefuture.org – The Monitoring the Future (MTF) survey is a series of classroom surveys of eighth, tenth, and twelfth graders conducted by researchers at the University of Michigan under a grant from the National Institute on Drug Abuse, part of the National Institutes of Health, US Department of Health and Human Services.

http://www.inhalants.org/guidelines.htm – The National Inhalant Prevention Coalition (NIPC).

http://teens.drugabuse.gov/facts/facts_inhale1.php – The National Institute on Drug Abuse (NIDA): information on inhalants.

http://www.oas.samhsa.gov/nhsda.htm – The National Survey on Drug Use and Health (NSDUH), formerly called the National Household Survey on Drug Abuse, is sponsored by the Substance Abuse and Mental Health Services Administration. The survey has been conducted since 1971 and serves as the primary source of information on the prevalence and incidence of illicit drug, alcohol, and tobacco use in the civilian, noninstitutionalized population aged 12 years or older in the United States.

http://www.whitehousedrugpolicy.gov/about/index.html – The White House Office of National Drug Control Policy (ONDCP) is a component of the Executive Office of the President and was established by the Anti-Drug Abuse Act of 1988.

CHAPTER

74

Ketamine

Jan Copeland, Paul Dillon

National Cannabis Prevention and Information Centre, University of New South Wales, Sydney, NSW, Australia

OUTLINE

INTRODUCTION

Ketamine is a short-acting, noncompetitive *N*-methyl-D-aspartate (NMDA) receptor antagonist that acts as a dissociative anesthetic with analgesic and amnestic properties. It has been used for humans and animals since the 1960s and its first reported use as a recreational drug was in the early 1980s. Both popular and research accounts indicate that the recreational use of ketamine has widened in the context of nightclubs, dance parties, and raves. Clinicians and researchers noted that some nonmedical users sought dissociative, hallucinatory experiences. By the end of the 1970s, the United States Food and Drug Administration (FDA) was becoming concerned about the sale of ketamine on the street. By the early 1980s a wide range of unauthorized preparations were available in the United States including capsules, powder, crystals, tablets, and solutions, in addition to the authorized injectable forms. Solutions sold on the street in the United States have gone by many names such as K, Kay, Jet, Super Acid, 1980 Acid, with powders known as Green, Purple, Mauve, Special LA Coke, Super C, and K. Other street names have included Vitamin K and most recently, Special K.

Ketamine is available as a liquid, pill, or powder, and can therefore be snorted, swallowed, smoked, or injected. It is typically snorted in measures known as bumps. Bumps are small snorts usually measured by

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00074-7

a tiny spoon provided within the lid of the container in which it is purchased.

HISTORY OF KETAMINE

Ketamine hydrochloride is marketed as a short-acting, general anesthetic for human and veterinary use. Reports have indicated that ketamine, or Special K as it is also known, is being used in social rather than medical and scientific settings in many parts of the world. It was first synthesized by Calvin Stevens in 1962. Early clinical studies on ketamine with human volunteers found it to be more effective and shorter acting than phencyclidine (PCP), with fewer emergence symptoms and less toxicity. The drug was first manufactured in the United States in the 1960s as Ketalar. It was described as a dissociative anesthetic with analgesic and amnesic actions. Use of ketamine as a surgical anesthetic escalated when it gained popularity on the battlefields of Vietnam. It was promoted as a dissociative anesthetic because of its ability to induce a lack of responsive awareness, not only to pain but also to the general environment. It is believed that the drug selectively interrupts association pathways of the brain before producing somesthetic (the consciousness of having a body) sensory blockade.

PHARMACOLOGICAL AND TOXICOLOGICAL EFFECTS OF KETAMINE

Ketamine is 2(2-chlorophenyl)-2-(methylamino)-cyclohexanone, an arylcycloalkylamine. It is structurally related to PCP and cyclohexamine. Ketamine is a noncompetitive NMDA receptor antagonist that interferes with the action of excitatory amino acids including glutamate and aspartate. It occurs in racemic form and also as an *S* enantiomer. Ketamine is manufactured by the chemical industry for use in pharmaceutical products using the precursors cyclopentyl bromide, *o*-chlorobenzontrile, and methylamine. Due to the complicated multistep synthesis, and the difficulty of obtaining the necessary precursors and numerous solvents and reagents, ketamine sold illicitly for recreational use appears to be mostly obtained by diversion of legitimate supplies of either the bulk drug or its pharmaceutical preparations.

Ketamine has a plasma half-life of 2–4 h. It is highly lipid soluble and has a distribution half-life of approximately 7–11 min. Ketamine is metabolized in the liver within the cytochrome P_{450} system with formation of water-soluble conjugates. Some of the metabolites, namely norketamine, have some potency but do not penetrate the central nervous system sufficiently to cause hypnosis. Ketamine is typically administered in doses of 1–2 mg kg^{-1} intravenously over 1–2 min or 4–5 mg kg^{-1} intramuscularly. Studies of the pharmacokinetics of ketamine report similar plasma concentration profiles for the same dose administered orally, sublingually, and in suppository preparations with nasal preparations having the highest relative concentration of the noninjecting routes of administration.

While ketamine has analgesic effects it is not reversed by naloxone. The synthetic alpha$_2$-adrenergic antagonist, atipamezole hydrochloride, is used in veterinary medicine to reverse the effects of medetomidine and ketamine anesthetics in a range of animals. There are no reports of its use in humans. At this stage, therefore, there is no agent available to reverse the effects of ketamine in humans.

PATTERNS OF USE

That ketamine-related data are not individually reported in the World Drug Report epidemiology of illicit drug use is testimony to its low level of use in general populations of most participating countries. It was not reported in the triennial Australian National Drug Strategy Household Survey until 2004. That year, 0.3% reported having used ketamine in the previous 12 months, compared with 0.2% in the 2007 survey. In the latter survey, the average age of initiation was 24.0 years (23.8 males; 24.3 females). The typical use pattern was once or twice a year (51.1%) but 6% used ketamine between daily and weekly. The other drugs most commonly reported as being used concurrently with ketamine were alcohol (73.5%), sildenafil citrate (Viagra®) (53.6%), and cannabis (24%).

In 1994, however, ketamine was listed in the Australian Illicit Drug Report for the first time. This is an Australian national annual survey of injecting drug users. An additional annual survey of regular ecstasy users has reported a decline between 2003 and 2009 from 26% of participants reporting having used ketamine in the previous six months in 2003 to only 10% in 2009.

The British Crime Survey shows that less than half of 1% of all respondents reported the use of ketamine in 2007–08, and that there was no statistically significant change in this figure compared to the previous year. Surveys of nightclub patrons in the United Kingdom, however, have reported an increase in ketamine use from 25 to 40% from 1999 to 2003. A study in 2009 found increased use in 9 out of 20 studied areas (London, Birmingham, Newcastle, Ipswich, Bristol, Blackpool, Portsmouth, Nottingham, and Sheffield) and that ketamine users were experimenting with stronger doses, including injecting the drug.

One of the few population level data collections to include ketamine is the US Monitoring the Future Study, an ongoing study of the behaviors, attitudes, and values of American secondary school students, college students, and young adults. The 2009 trends for 12th graders reported a reduction in annual ketamine use from a peak of 2.6% in 2002 to 1.7% in 2009.

The use of ketamine in Asia appears to be relatively more common. A Taiwanese study reported that ketamine was detected in 47% of the urine samples obtained from a study of those attending rave parties but in only 2% of the urine of police detainees. A recent study of drugged drivers involved in nonfatal driver casualties in Hong Kong reported that ketamine was the most commonly detected substance, being found in 45% of participants. In Shanghai and neighboring cities, ketamine was more frequently detected (0.03%) than 3,4-methylenedioxymethamphetamine (MDMA) (0.1%) in the serum of drivers involved in motor vehicle accidents and traffic violations.

The most recent Drug Abuse Warning Network report on trends in drug-related emergency department visits for 1994–2001 reports a 3474% increase in the number of drug mentions for ketamine from 19 in 1994 to 679 in 2001. More recently, from 2005 to 2008, the number has stabilized to between 300 and 344 ketamine mentions per year.

EFFECTS OF KETAMINE

Ketamine has systemic effects on a number of organ systems. The predominant effect of ketamine on the cardiovascular system is thought to be due to decreased catecholamine reuptake leading to increase in arterial blood pressure and cardiac output 2–4 min after intravenous injection and 10–20 min after intramuscular injection. A recent study has reported that ketamine reduces left ventricular systolic and diastolic functions among those with ischemic heart disease.

Ketamine may cause an increase in cerebral blood flow, oxygen consumption, and intracranial pressure. Animal studies, however, have demonstrated a marked neuroprotective effect, mediated by antagonism of the NMDA channels on central neurons, and may have promise in the future management of cerebrovascular accidents. Upper airway reflexes are usually maintained when ketamine is administered in addition to bronchodilation. Apnea has been seen with ketamine use but is generally associated with rapid and/or large intravenous doses. Additional effects are seen in the gastrointestinal tract (increased salivation), the immune system (a significant reduction in leukocyte activation during sepsis), and the eyes (initial rise in intraocular pressure, eye movements throughout surgery, and preservation of the corneal reflex).

Key Psychoactive Effects of Ketamine

In one of the earliest studies of the recreational use of ketamine, it was reported that the drug was viewed by most of the users in the study as a safe, potent hallucinogen with a short duration of action and an equal balance between positive and negative effects. Self-administration was titrated to achieve the desired amount of dissociative sensations, hallucinations, and transcendental experiences. Respondents reported ataxia, slurring of speech, dizziness, mental confusion, blurred vision, anxiety, hyperexcitability, and insomnia, amongst other effects.

In a later study that examined the subjective effects of ketamine, seven male volunteers were given between 5 and 12 doses of ketamine intermittently over a period of 18 months. They used approximately 10–25% of the anesthetic dose, believing this to reflect the dose level that occurs during the emergence period. The authors describe a series of phenomena experienced by most of the participants: a sensation of light throughout the body; novel experiences concerning body consistency (e.g. feeling as though they are made from wood or plastic); grotesquely distorted shape or size of body parts; sensation of being weightless and floating or hovering; colorful visions (including geometric patterns and figures); absence of sense of time; sudden insights into the nature of the existence or the self; strong feelings of association with others in the environment; and out-of-body experiences.

Subsequent studies of the psychedelic effects of ketamine in healthy volunteers have reported that these effects are dose related. Anecdotal evidence indicates that many users of other club drugs have hesitated to use it, due to either bad personal experiences or horror stories relayed to them by others. Some users may describe visits to the K hole or K ride as it is known in Asia, as a place referring to where users are when under the influence of ketamine. The K-hole experience appears to vary with the individual, but six main categories of mental effects produced by ketamine have been identified: the perception of contact with aliens; the perception of entry into information networks; access into alternative realities; personal and creative problem solving; out-of-body near-death states; and tantra-like enhancement of sexual activity.

Ketamine can sometimes reproduce the features of a near-death experience (NDE), including buzzing/ringing/whistling sounds at the beginning, travel through a dark tunnel into light at a high speed, the conviction that one is dead, apparent telepathic communion with God, intense visions, and out-of-body

experiences. A survey of 50 ketamine users who reported experiencing an NDE reported that it typically occurred in their first few occasions of use with the most frequently noted features being altered perception of time (90%), strong sense of detaching from their own body (88%), and a sense of peace and joy (76%). It should be noted that an NDE induced by ketamine does not necessarily mean that the person is physically near death.

Psychosis

The role of the NMDA neurotransmitter system in relation to psychosis is not completely understood, however, pharmacological, postmortem, and clinical studies have implicated the NMDA system has been implicated in the pathophysiology of schizophrenia. The psychotic states induced by NMDA antagonists such as ketamine have been described as being similar to schizophrenia. The hallucinogenic effect of ketamine arises, at least in part, from its capacity to disrupt thalamocortical gating of external and internal information to the cortex. Deficient gating of sensory and cognitive information is thought to result in an overloading inundation of information and subsequent cognitive fragmentation and psychosis. Animal studies have shown that treatment with ketamine leads to an increase in D2 receptor binding in the hippocampus and a decrease in glutamate receptor binding in the frontal cortex with no change in D1 receptor binding. The density of dopamine receptors was increased in the striatum and 5-hydroxytryptamine transporters were increased in the striatum, the hippocampus, and the frontal cortex.

Laboratory studies of subanesthetic doses of ketamine in humans report that it produces behaviors similar to the positive (e.g. hallucinations and thought disorder) and negative (e.g. detachment, amotivation, and blunted affect) symptoms of schizophrenia, elicits alterations in perception, impairs performance on tests of vigilance, verbal fluency, and the Wisconsin Card Sorting Test, evokes symptoms similar to dissociative states, and preferentially disrupts measures of frontal lobe function such as delayed word recall, sparing immediate recall, and postdistraction recall. Ketamine also exacerbates the core psychotic symptoms in patients with schizophrenia and is not blocked by antipsychotic drugs such as haloperidol.

As a result ketamine is a useful model of schizophrenia and is being used to increase our understanding of the illness and to develop new treatments.

Depression

A growing body of preclinical research implicates the NMDA class of glutamate receptors in the pathophysiology of major depression and the mechanism of action of antidepressant treatments. A small study of those with diagnosed major depression reported robust decrease in depressive symptoms, emerging progressively over 3 days, following low-dose ketamine infusion. There are studies currently being conducted in the United States on the use of ketamine as treatment for bipolar disorders.

Teratogenic Effects

Ketamine crosses the placenta easily and concentrations in the fetus are approximately equal to those of the mother. There is accumulating evidence from in vitro and in vivo experiments that suggests that tonic stimulation of NMDA receptors is vital for the survival of developing nerve cells. The NMDA receptor can be pharmacologically blocked by drugs such as alcohol and ketamine, among other drugs. This alters the glutamate and γ-aminobutyric acid (GABA) transmission that suppress neuronal activity and cause neuronal death in the developing brain. While there are no human studies, animal studies suggest that the duration of ketamine exposure in 7-day-old rat pups increases neuronal degeneration of the developing rat brain.

Safety Profile

Ketamine has a wide margin of safety. The predominant adverse reaction is emergence phenomena post anesthetic for which the risk factors are age, female gender, noisy environment during recovery, prior personality disorders, and excessive dreaming. The experience of emergence phenomena is not associated with dose. Other side effects include a transient rash predominantly on the face and neck, and nausea and vomiting. A reasonably conservative safety ratio of 35 for snorted ketamine and a somewhat higher margin for oral administration has been reported.

THERAPEUTIC VALUE OF KETAMINE IN HUMANS

Ketamine has broad areas of application and is a rapidly acting, relatively safe parenteral analgesic and anesthetic agent that has been in clinical use for more than three decades. It is the only injectable anesthetic that induces increase in arterial pressure and heart rate. Although the respiratory and pharyngeal reflexes are sometimes momentarily depressed after injection of substantial quantities, they are usually maintained during the period of unconsciousness. The drug is therefore suitable for short anesthetic and surgical procedures especially in the absence of a trained anesthetist, although the latest Parke-Davis data sheet stresses that

a trained professional should be present, together with resuscitation equipment. It is particularly useful in developing countries and remote country areas where a doctor may be working alone.

The major concern is the emergence phenomena. Psychic disturbances after ketamine anesthesia have been reported to occur in about 15–40% of adult cases, depending to some extent on how these terms are defined. Other drugs, such as diazepam, lorazepam, and propofol, have been given together with ketamine in an attempt to reduce or abolish these phenomena, with some success. Psychological techniques are also effective in reducing complaints. The emergence phenomena have led to less medical use than was originally anticipated, but ketamine is still to be found in many general hospitals in most countries. It is also currently widely used in veterinary medicine. A number of clinical uses of ketamine have been suggested. Ketamine appears to be best used in the young (less than 10 years old) and the old (over 60 years) as these groups have reported fewer emergence reactions (Radford, 1996). Recent studies have confirmed the effectiveness of ketamine in a variety of pediatric and diagnostic procedures via caudal or intramuscular routes of administration.

In addition to anesthesia and sedation, ketamine is used as an analgesic for acute postoperative pain, painful procedures, and more controversially for chronic or neuropathic pain. Analgesic doses of ketamine have been associated with dose-related declines in mood, conscious perception, and intellectual performance. Recent studies also report the successful use of patient controlled analgesia with morphine and ketamine following spinal and hip surgery in adults. An early stage descriptive study of the use of topical ketamine in the management of postherpetic neuralgia suggests it shows promise in the management of this painful condition.

Ketamine has also been used as an experimental treatment for complex regional pain syndrome, which is a severe, chronic pain condition characterized by sensory, autonomic, motor, and dystrophic signs and symptoms. The treatment involves one or two cycles of a 10-day, inpatient intravenous infusion for 4 h daily in high doses. While there have been promising findings, the studies to date have been open label and the evidence base is yet to be developed.

There is mixed evidence on the effect of ketamine on epilepsy and seizure disorders. Ketamine has been reported as both a pro- and anticonvulsant. Recently, ketamine has been recommended for use with electroconvulsive therapy as it has been shown to prolong seizure duration with more rapid posttreatment reorientation. A larger body of evidence, however, suggests that it displays anticonvulsant and neuroprotective properties but should be used with caution over prolonged periods of time.

Studies conducted in the 1950s and 1960s suggested that psychedelic drug-assisted psychotherapy might be an effective treatment for alcohol dependence. In the 1990s, these findings were used as a partial basis for ketamine-assisted treatment of alcohol dependence. This is known as ketamine psychedelic therapy (KPT). The same author reports on the use of ketamine psychotherapy for heroin addiction, where dose was found to be related to treatment outcome at 2 years. These are controversial treatments that are yet to be used outside Russia.

In the current environment of growing concern about bioterrorism, the neuroprotective and antiepileptic activities of ketamine have led to suggestions that it may be a useful tool in the management of nerve agent poisoning such as sarin.

KETAMINE-RELATED MORTALITY

There are very few deaths by pure ketamine overdose recorded (i.e. not also involving a drug such as alcohol). Of 87 ketamine-linked deaths in New York City, none was purely due to the use of ketamine. Parke-Davis have reported that there are cases of accidental injections with ten times the amount required for surgery, with no obvious lasting effects. The principal physical dangers of most nonmedical uses are currently believed to arise mainly from the setting, or an interaction between the user and the setting of use, as ketamine can leave the user in a confused state. This can, for example, result in burns, falls (sometimes fatal), drowning, death by hypothermia from lying outside in winter, traffic accidents, and becoming a crime victim (e.g. drug rape). A recent paper reported the case of an emergency medical technician who died of a combination of asphyxia and intoxication with administered intravenous ketamine in an autoerotic accident.

There are two reports in the literature of deaths by pure ketamine overdose. One described as a homicide for homosexual ends. The other described a middle-aged woman who took the drug daily for 7 months.

KETAMINE-RELATED MORBIDITY

There is an extensive list of possible physical effects of ketamine that may be seen as adverse by the user, or that may be directly harmful. Some of those of principal concern in a nonmedical use context are difficulty with walking and balance resulting in falls, numbness, slurred speech, dizziness, visual problems, nausea, headaches, spasms, and twitches. A recent study found

that, consistent with the animal studies, frequent ketamine use produces long-lasting impairments in episodic memory and aspects of retrieval from semantic memory even when other drug use is taken into account. A later study reported that when retested 3–4 years later following significant reductions in levels of ketamine use, semantic memory impairments associated with ketamine use were reversible. The effects upon schizotypal symptoms and perceptual distortions may, however, persist.

The use of ketamine has been linked with a range of unpleasant mental effects including anxiety, panic attacks, flashbacks, posttraumatic stress disorder, persistent perceptual changes, mania, depression, suicide, insomnia, nightmares, night terrors, an unpleasant feeling of being unreal or that the world is unreal, paranoid delusions, persistent hallucinations, automatic behavior, and fragmentation of the personality and aggression.

A study of 23 recreational users noted a high incidence of flashbacks and attentional dysfunction, but exactly what was meant by flashbacks is not defined. Large anesthetic studies do not confirm the finding and generally conclude that ketamine is usually devoid of significant persistent effects once the drug and its metabolites have cleared the body. For example, in a study of 1400 patients given ketamine as an anesthetic for surgical procedures, three had prolonged hallucinations, none lasting beyond 3 weeks. In no case did hallucinations begin after a period of normality, which is integral to the World Health Organisation (1992) definition of flashbacks.

A 2009 study of the consequences of chronic self-administration on the neurocognitive function and psychological well-being involved 150 individuals. The study over a period of 1 year compared levels of ketamine use frequency (frequent ketamine users, infrequent ketamine users) with abstinent users, polydrug using controls, and drug-free controls. Among frequent ketamine users, increasing use was correlated with decreasing cognitive performance as measured by spatial working memory and pattern recognition memory tasks. Assessments of psychological well-being also showed greater dissociative symptoms among frequent users and a dose–response effect on delusional symptoms.

Until recently, there had only been reports that ketamine could cause toxic changes in the rat brain. A Chinese study of 41 ketamine-dependent individuals compared with 44 healthy controls found white matter changes associated with ketamine use in the bilateral frontal and left temperoparietal cortices. These changes were correlated with the severity of ketamine use. These findings suggest a microstructural basis for the changes in cognition and experience observed among those using ketamine for prolonged periods. The similarities of these changes to those observed in chronic schizophrenia, moreover, have implications for the glutamate model of the illness. The use of ketamine with other neurotoxic drugs like alcohol is an additional cause for concern.

Urological and Upper Gastrointestinal Effects

A number of series of case reports have recently identified physical health consequences of ketamine use. A review of 233 ketamine-related presentations to 15 Hong Kong emergency departments reported that the most common presenting symptoms were impaired consciousness (45%), abdominal pain (21%), lower urinary tract symptoms (12%), and dizziness (12%).

A retrospective study of 64 heavy ketamine users explored the effects on the upper gastrointestinal tract. When those with other gastrointestinal conditions were excluded, 37 participants were studied. Of these, 28 (75.7%) reported upper gastrointestinal symptoms including epigastric pain and vomiting. It was further reported that despite medications, symptoms tended to persist unless ketamine use ceased.

A more concerning consequence of ketamine use that has been newly identified is so-called ketamine bladder syndrome, which was first described in 2007. Clinicians from Canada described nine patients who were daily ketamine users and who presented with severe dysuria, frequency, urgency, and gross hematuria. At cystoscopy, all patients had severe ulcerative cystitis. Clinicians in Hong Kong reported a similar presentation of ten street ketamine users diagnosed with ulcerative cystitis.

This unwanted effect of ketamine has also been reported among pediatric patients in less than 2 weeks after commencing use for chronic pain. The largest study to date was conducted in Hong Kong and reports on 59 ketamine users with moderate to severe lower urinary tract symptoms. Patients had frequency rates from 15 to 90 min between voids. Functional voiding capacity was between 20 and 200 ml, which represents a dramatic decrease in bladder capacity. Upon cystoscopy, 71% of these patients showed various degrees of bladder mucosal inflammation similar to that seen in interstitial cystitis. Histological examination showed a largely denuded bladder epithelium and granulation of the lamina propria infiltrated by lymphocytes and eosinophils. Most patients showed detrusor over activity at very low bladder infusion volumes. Even more concerning is the documentation of severe kidney effects among these individuals, where 51% showed unilateral or bilateral hydronephrosis on renal ultrasonography and 7% had features suggestive of papillary necrosis. Similar case series are now being reported in the United Kingdom and other regions of China.

The urologists treating this disabling condition have also reported on the apparent difficulty their patients experienced in abstaining from ketamine use. A group in Taipei reported that of ten consecutive patients with ketamine bladder syndrome, where clear advice and multidisciplinary support was provided to cease their ketamine use, at 12 months follow-up only 3 had been able to do so successfully.

KETAMINE DEPENDENCE

There is a substantial amount of popular literature describing ketamine as having a marked potential for giving rise to nonphysical dependence and case studies in the medical literature are accumulating. Many of these early reports of ketamine dependence were among those with access to the drug such as anesthetists and veterinarians.

The Australian National Minimum Dataset of Clients of Specialist Drug and Alcohol Treatment Services records no identified episodes of care for ketamine as the principal drug of concern in its collection over the past decade. The comparable dataset in the United States reported a total of 229 mentions in 2006 where in only 82 was ketamine the primary drug of concern. While this appears to indicate that ketamine is either used by only a small number of individuals or that it is unlikely to cause dependence sufficient to drive treatment seeking, it may be that, as there is no evidence-based intervention for ketamine-related problems, that they are not admitted to treatment services.

There is evidence from animal studies to support the view that ketamine can give rise to a dependence syndrome without physical withdrawal phenomena. This resembles cocaine dependence without the crash after use. Ketamine is also self-administered in rats and nonhuman primates.

A recent Australian study of 100 recreational ketamine users found that around one in five (22%) of participants reported physical tolerance to ketamine. This is consistent with reports of tolerance to ketamine following multiple anesthetics.

There have been at least two reported case histories of inpatient withdrawal from ketamine with observed psychotic symptoms. One case also included reports of a number of previously experienced ketamine withdrawal symptoms such as chills, autonomic arousal, lacrimation, restlessness, nightmares, and psychological craving with ketamine being used to relieve the symptoms.

There is no research literature on the management of ketamine abuse or dependence. Case histories of ketamine withdrawal report the use of benzodiazepines for the management of anxiety and insomnia and vitamin supplements. There are some generic descriptions of potentially useful interventions that suggest an abstinence-oriented approach similar to that used for psychostimulants.

PUBLIC HEALTH ISSUES ASSOCIATED WITH KETAMINE USE

In 1980, nonmedical unauthorized experimentation with ketamine was first reported in Australia. The authors believed that use of the drug was largely confined to medical circles. Since that time, anecdotal evidence indicates that the nature of users has become much more varied.

In Australia, ketamine is often marketed in small glass vials with small spoons to measure accurate doses. The spoon is contained in the cap of the bottle, much like a snuff spoon. Other bottles have a self-contained measurer in the cap, which when turned over leaves a measured amount, which can then be snorted as bumps; very few ketamine users snort the drug in lines. This may result in lower doses per snort than its use in lines as is seen more commonly internationally. The effects of differing methods of insufflation on health outcomes have not been examined.

Ketamine appears to be obtained from the diversion of legitimate supplies (veterinarians and pharmaceutical companies) or is imported from overseas. Ketamine can be bought over the counter in some Asian countries. Ketamine sold illicitly is often converted from a liquid form to a powder utilizing a simple evaporation process. The liquid ketamine is dried in a variety of ways (microwave oven or sun-dried) until a residue remains. This crystal residue is ground into a powder, leaving a fine powdery material similar to cocaine and heroin. In this form, it is far more convenient and more marketable than the injectable drug.

Evidence from the United States indicates that, up until 1996, ketamine was usually not adulterated with other substances. Of all the ketamine submissions to the US Drug Enforcement Agency (DEA) regional laboratories in that period, there was only one instance of this type of dilution recorded. However, this situation may well have changed when price and demand rose sufficiently to make such practices profitable.

Both popular and research accounts indicate that the recreational use of ketamine has widened in the context of nightclubs, dance parties, and raves. This has caused concern as ketamine is an anesthetic. One study concluded that it was totally inappropriate to use ketamine as a dance drug. The reasons for this were factors such as set and setting, the rapid onset of the drug, and the intensity of the experience as a whole. The respondents believed that using ketamine in a noisy, busy, or crowded environment was potentially dangerous and

that use should be confined to a familiar and secure place, such as one's home. All users had been unprepared for the intensity and nature of the effects when they first used the drug. In a study of the toxic effects of club drugs, it was concluded that the "single most risk-producing behavior of club drug users is combining psychoactive substances, usually involving alcohol."

The use of ketamine and other party or club drugs is reportedly associated with an increased incidence of unsafe sex among gay men on the circuit party scene in the United States. This is clearly a public health issue in the spread of blood-borne viruses and other sexually transmitted diseases. A small study of high-risk youth injecting ketamine in New York City also reported that its use was associated with a range of high-risk injection practices such as group injection with shared paraphernalia. Recent surveys among this sentinel group, however, have shown declines in the levels of ketamine (and other drug use) between 2002 and 2007.

An Australian study reported a number of ketamine-related problems among their sample of nonmedical ketamine users. While only one in five stated that they had ever experienced severe side effects as a result of ketamine use, more than one-third (38%) reported having to deal with someone else who had suffered badly following ketamine use. More than half (58%) of those interviewed had experienced the K hole and this was related to increased exposure to the drug – having used more than 20 times. The most commonly reported problems were employment related (20%). These included vagueness affecting work performance and lesser volume of work being produced. Additional problems reported included relationships (5%), financial (5%), and legal (1%). Of the five participants reporting relationship problems, none also reported financial problems but three of the five also reported work-related problems. Twenty-two participants reported at least one problem, five reported two problem areas, and one reported three problems areas. Ketamine has been sold in the rave scene in the United Kingdom as a key component in fake MDMA (ecstasy) tablets. There are also reports of ketamine being sold as ecstasy in Australia, or used as a cutting agent in other drugs such as cocaine, amphetamines, and heroin.

THE LEGAL STATUS OF KETAMINE

Ketamine is scheduled in many jurisdictions. Below are some examples of the legal status of ketamine internationally.

Australia

Ketamine is scheduled differently within and between countries. Below is a brief overview of the legal status of ketamine in a number of key jurisdictions. Within Australia prior to December 2003 ketamine was a scheduled drug in New South Wales under the Poisons and Therapeutic Goods Act, which meant that there were strict regulations on the sale and distribution of the substance. In response to increasing trends in ketamine use, the government added it to the list of prohibited drugs in New South Wales. This means that it has now been listed under the Drug Misuse and Trafficking Act 1985, ensuring that tougher penalties are in place for dealing with its illicit use. New penalties for those caught manufacturing or supplying the drug include fines between AU$5500 and AU$550 000 and anywhere from 2 years to life imprisonment or both. These penalties particularly apply to veterinary and medical practitioners who have an authority to purchase ketamine for legitimate purposes and then supply it to others for illicit use. Similar laws apply in some other states of Australia, however, in other states it remains legal to possess ketamine.

Canada

Ketamine is Schedule 1 in Canada, pursuant to a February 2004 notice that it "is an analog of phencyclidine (PCP), and is, therefore, captured as item 14 in Schedule I of the CDSA and item 14 in the Narcotic Control Regulations (NCR)." Affected researchers were expected to have complied with the Schedule 1 protocol for the handling of ketamine by August 31, 2005.

China

Ketamine is a Class II psychiatric drug; trade is limited to licensed wholesalers and retail sales are prohibited. In July 2004, the state of Sichuan placed ketamine in Class I; other provinces are expected to follow suit.

Mexico

Ketamine is a Category 3 drug under Mexico's General Health Law; Category 3 drugs "have a therapeutic value but constitute a problem for the public health." Medicines Administration Regulations restricts the acquisition of ketamine to licensed veterinarians only and sets strict rules regarding the management and follow-up of the products.

Netherlands

Ketamine is treated as a medication and is not in List 1 or 2 in the Netherlands.

United Kingdom

Prior to January 1, 2006, ketamine was not controlled in the United Kingdom under the Misuse of Drugs Act, making it legal to possess. However, sales and distribution were controlled under the Medicines Act making it illegal to sell or distribute without a license. Ketamine became a Class C drug on January 1, 2006, following a 2004 report to government that recommended moving ketamine into Class C.

United States

Ketamine was unscheduled until August 1999. In early July 1999, the DEA added ketamine to Schedule III with an emergency ruling that took effect on August 12, 1999. It is now a federal offense to possess ketamine in the United States without a license or prescription. Since that time, ketamine has been scheduled in many individual states, and laws are pending in many more. Where ketamine is not scheduled in a specific state, all prosecutions for possession or sales occur at the federal level.

SUMMARY

In conclusion, as levels of ketamine in the general population appear to be very low, the harms reported by recreational users are not excessive, adulteration is rare, and the mortality rate is low. Ketamine does not appear to pose a risk to public health at this time. At the individual level, many of those who experiment find the effects aversive and do not persist. The harms that require further investigation are the association with unsafe sex and injecting behaviors, the neurotoxic effects, and use in situations where there is a heightened risk of accidental death when the user's cognition is grossly impaired. It is recommended that the usual harm minimization strategies; especially not to use the drug when alone, not to concurrently use other neurotoxins such as alcohol, to be in a physically safe environment, and to use safer injecting techniques; should be observed when using ketamine.

List of Abbreviations

NDE near-death experience
NMDA *N*-methyl-D-aspartate
PCP phencyclidine

Glossary

Bumps bumps are a measure of ketamine that is a small snort usually measured by a tiny spoon provided within the lid of the container in which it is purchased.

K hole this is a slang term for a state of dissociation from the body that may mimic the phenomenology of schizophrenia. The K-hole experience appears to vary with the individual and includes intense hallucinations and perceptual distortions.

Further Reading

Adler, C.M., Malhotra, A.K., Elma, I., Goldberg, T., Egan, M., Pickar, D., et al., 1999. Comparison of ketamine-induced thought disorder in healthy volunteers and thought disorder in schizophrenia. American Journal of Psychiatry 156 (10), 1646–1649.

Bowdle, T.A., Radant, A.D., Cowley, D.S., et al., 1998. Psychedelic effects of ketamine in healthy volunteers. Anesthesiology 88 (1), 82–88.

Curran, H.V., Monaghan, L., 2001. In and out of the K-hole: a comparison of the acute and residual effects of ketamine in frequent and infrequent ketamine users. Addiction 96, 749–760.

Curran, H.V., Morgan, C., 2000. Cognitive, dissociative and psychogenic effects of ketamine in recreational users on the night of drug use and 3 days later. Addiction 95 (4), 575–590.

Dillon, P., Copeland, J., Jansen, K., 2003. Patterns of use and harms associated with non-medical ketamine use. Drug and Alcohol Dependence 69, 23–28.

Dotson, J.W., Ackerman, D.L., West, L.J., 1995. Ketamine abuse. Journal of Drug Issues 25 (4), 751–757.

Jansen, K.L.R., 2001. Ketamine, Dreams and Realities. Multidisciplinary Association for Psychedelic Studies, Florida.

Morgan, C.J.A., Monaghan, L., Curran, H.V., 2004. Beyond the K-hole: a 3-year longitudinal investigation of the cognitive and subjective effects of ketamine in recreational users who have substantially reduced their use of the drug. Addiction 99 (11), 1450–1461.

Vollenweider, F.X., Kometer, M., 2010. The neurobiology of psychedelic drugs: implications for the treatment of mood disorders. Nature Reviews Neuroscience 11, 642–651.

Weiner, A.L., Vieira, L., McKay, C.A., Bayer, M.J., 2000. Ketamine abusers presenting to the emergency department: a case series. Journal of Emergency Medicine 18 (4), 447–451.

White, J.M., Ryan, C.F., 1996. Pharmacological properties of ketamine. Drug and Alcohol Review 15 (2), 145–155.

CHAPTER

75

Anabolic-Androgenic Steroid Use and Dependence

Gen Kanayama, Harrison G. Pope, Jr.
McLean Hospital, Belmont, MA, USA

OUTLINE

DEFINITION OF ANABOLIC-ANDROGENIC STEROIDS

The anabolic-androgenic steroids (AAS) are a family of hormones that includes testosterone, the naturally occurring male hormone, together with numerous synthetic derivatives of testosterone that have been created over the last 70 years. As their name implies, AAS have both anabolic (muscle-building) properties and androgenic (masculinizing) properties; there are no AAS that can produce one of these effects without the other. AAS must not be confused with other types of steroids, such as corticosteroids. Corticosteroids are a separate family of hormones that includes the natural hormone cortisol, together with many synthetic derivatives of cortisol; corticosteroids have a similar chemical structure to AAS, but very different biological actions. In particular, corticosteroids do not lead to muscle gain, and thus are not used by athletes or by other illicit drug abusers.

HISTORY

As early as the late nineteenth century, the noted physician Brown-Séquard injected himself with an extract that he had prepared from testicles of guinea pigs and dogs, and noted a subjective improvement in energy and virility. However, it appears that his preparation actually contained no active ingredients. It was not until the 1930s that testosterone was ultimately synthesized by chemists in Germany (see Fig. 75.1 for a chronology of this and subsequent events). Over the next decade, dozens of synthetic derivatives of testosterone were created, yielding the family that we now call AAS. It is rumored that Hitler gave AAS to some of his troops during the 1940s in order to make them more aggressive, although no written documentation of this story seems to exist.

Also in the 1940s, physicians began to prescribe testosterone and other AAS to male patients, especially in attempts to treat symptoms of depression in aging

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00075-9

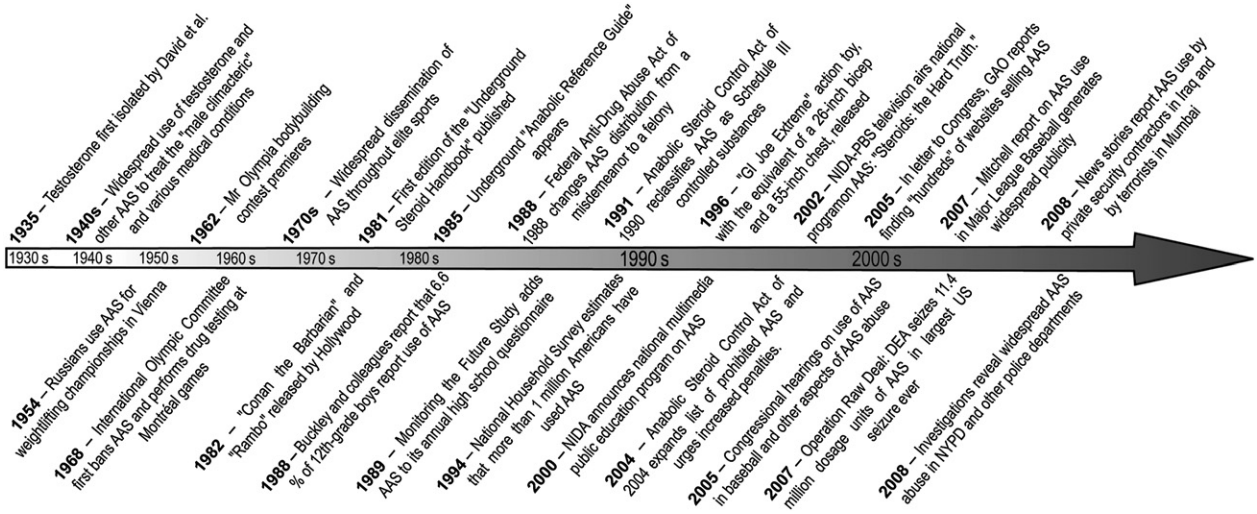


FIGURE 75.1 A timeline illustrating various developments in the chronology of AAS use. *Adapted from Kanayama, G., Hudson, J.I. and Pope, Jr., H.G. (2008). Long-term psychiatric and medical consequences of anabolic-androgenic steroid abuse: a looming public health concern?* Drug and Alcohol Dependence *98, 1–12.*

men. A substantial literature appeared in that decade, reporting considerable success in using testosterone and its relatives as antidepressant medications. One study even reported that the efficacy of testosterone in depression was comparable to that of electroconvulsive therapy. However, by the early 1950s, the discovery of other antidepressant medications, such as the tricyclic antidepressants and monoamine oxidase inhibitors, eclipsed the use of testosterone as an antidepressant, and there were few additional publications on this topic over the next several decades.

By the 1950s, athletes had begun to discover the remarkable efficacy of AAS for sports, including particularly sports involving muscle mass and strength, such as weightlifting, power lifting, and field events such as the shot put. Perhaps the first well-documented case of AAS use was by the Russian weightlifting team at the world championships in Vienna, Austria, in 1954. By the 1960s, AAS use was widespread throughout elite athletes around the world, and by the late 1960s, AAS were officially banned in the Olympics. Formal drug testing for AAS was first instituted at the Olympic games in Montréal in 1968.

By the 1960s, AAS also became virtually mandatory for professional bodybuilders, since these drugs permitted bodybuilders to attain a degree of muscle mass far beyond that attainable by any natural human being. Bodybuilding shows and magazines began to proliferate; the Mr Olympia contest was initiated in 1962, and many other bodybuilding competitions followed. Still, most people in the general public were not aware of the extraordinary anabolic potency of AAS, and as recently as 1977, the American College of Sports Medicine published a position paper stating that AAS were not effective for muscle gains – a statement that was retracted 10 years later.

Through the end of the 1970s, AAS use remained largely in the domain of elite competitive athletes. Then, starting in about 1980 in the United States, AAS began to percolate out of the athletic world and onto the street. Larger and larger numbers of rank-and-file weightlifters, many of whom had no competitive athletic aspirations at all, began to discover AAS and to use these drugs. This evolution was fueled in part by the appearance of various underground guides offering information about how to use AAS, how to perform injections, and what results to expect. The first of these guides, the *Underground steroid handbook*, appeared in 1981, and quickly went through increasingly expanded revised editions. The 1980s also witnessed a surge in Western media images of supermuscular males, appearing in comic strips, advertisements, television dramas, and Hollywood movies such as *Rambo* and *Conan the Barbarian*. Even boys' action toys, such as GI Joe in the United States and Action Man in the British Commonwealth, have become steadily more muscular over the last several decades.

By the 1980s, epidemiologic studies in the United States began to appear, reporting AAS use in 3–11% of male high school students (see further details below). Since that time, illicit AAS use has continued to spread through young American men up to the present. Even using conservative estimates, at least 2 million American men have now used illicit AAS at some time in their lives. In European countries, widespread illicit AAS use began to appear about a decade behind the United

States; since the 1990s, studies in the United Kingdom, Scandinavia, and more recently elsewhere in Europe have documented lifetime AAS use in 2–6% of boys and young men.

Concern about AAS use has led to increasing legislation to try to prevent the spread of AAS, especially in the United States. In 1989, the United States Congress enacted laws making it a felony to sell AAS, and in 1991, Congress passed the Steroid Trafficking Act, placing AAS under the jurisdiction of the United States Drug Enforcement Administration, with increased penalties even for possession of AAS. Within a short time after 1991, domestic production of AAS within the United States was largely eliminated. Currently the legal status of AAS in other countries is variable; for example, in the United Kingdom it remains legal to possess a personal supply of AAS, whereas sale of AAS is illegal. In many other European countries both possession and sale of AAS are illegal. Despite regulation, however, illicit AAS use continues to flourish, because with the rise of the Internet, AAS can be purchased from countries where they remain legal, and interdiction of such shipments is difficult. Thus, it seems likely that illicit AAS use will remain a widespread form of substance abuse for the foreseeable future.

Returning to legitimate medical uses of AAS, it is notable that the therapeutic use of testosterone has enjoyed a renaissance of interest since the 1990s. Several factors appear to account for this change. First, it has been found that testosterone and other AAS are effective for treating the wasting syndrome associated with human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS), and testosterone is now widely prescribed for this purpose. Some controlled studies also suggest that testosterone exerts antidepressant effects in individuals with HIV/AIDS. Second, testosterone has now been marketed in the form of patches and gels that allow the drug to be administered without the necessity for injections (which had previously been required because testosterone is ineffective when taken by oral administration). The appearance of these formulations has made testosterone administration more user friendly, thus likely stimulating more widespread prescription. Third, there has been increased psychiatric interest in the antidepressant properties of testosterone, with several studies suggesting that testosterone may be of value in men with dysthymia or major depression – an observation harking back to the early reports of the 1940s described above. In recent decades, there has been increasing interest in the use of testosterone and other AAS to counteract the natural decline in testosterone that occurs in aging men. The merits of this treatment continue to be debated, with cautious medical practitioners at one extreme noting that the long-term hazards of testosterone replacement remain poorly understood, and with sometimes questionable antiaging specialists at the other extreme offering testosterone and other AAS to virtually any customer, even in the absence of clear medical indications.

In summary, both the licit and illicit use of AAS have expanded substantially in the last few decades. The remainder of this chapter focuses almost entirely on illicit AAS use, which represents one of the newest and most rapidly evolving major forms of substance abuse worldwide.

PREVALENCE OF ILLICIT AAS USE

The great majority of illicit AAS users are young men, mostly ranging in age from 18 to 40 years. Girls and women rarely use illicit AAS, since women typically do not desire to become highly muscular, and women are also vulnerable to the masculinizing effects of AAS, such as deepening of the voice, beard growth, and masculinization of secondary sexual characteristics. Not surprisingly, therefore, very few published studies have actually recruited and personally evaluated any female AAS users, and even the largest such study recruited only 25 such cases. Surprisingly, however, many anonymous surveys of high school students have quoted seemingly high rates of AAS use among girls as young as 12 or 13 years old. Upon inspection, it appears that these studies likely have produced large numbers of false-positive responses because of ambiguously worded questions about steroid use. For example, anonymous questionnaires may use a question such as, "have you ever used steroid pills or shots without a doctor's prescription?" However, this question does not distinguish between illegal AAS used for muscle growth, as opposed to corticosteroids that are widely used for legitimate medical purposes, or over-the-counter supplements from health food stores that a student might erroneously believe are steroids. Thus, a young girl may answer "yes" to the steroid question when in fact she has simply used hydrocortisone cream for poison ivy, or corticosteroids in an asthma inhaler, or because she has used a nutritional supplement that she erroneously believes was a steroid. A recently published analysis of anonymous survey questions, covering studies performed in the United States, Europe, Australia, and Brazil, has concluded that the true rate of genuine AAS among girls is likely less than 0.5%.

Among young boys below age 18 years, it is likely that AAS use has also been overestimated by many anonymous surveys, although there is no doubt that

some boys of high school age do in fact use genuine AAS either for athletic purposes or for improving personal appearance. After correcting for possible false-positive responses on anonymous questionnaires, it seems likely that the true rate of AAS use in boys below the age of 18 years is about 1–3%. However, a number of studies suggest that the onset of AAS use in men occurs after age 18 years, with a peak in the early 20s. Thus the great majority of AAS users are men between about 20 and 40 years of age. It follows that among men who have reached their late 20s, the prevalence of lifetime AAS use may be considerably higher than 1–3%. At the least, it seems reasonable to estimate that there are at least 10 million men – primarily in the United States, Scandinavia, British Commonwealth countries, other European countries, and Brazil – who have used AAS at some time in their lives. Importantly, because of the fact that illicit AAS use did not become widespread until after 1980, it follows that the oldest members of this illicit AAS-using population – those who first initiated AAS use as youths in the 1980s – are only now reaching middle age. These middle-aged men, comprising both former AAS users and some still-current users, represent effectively the leading edge of an epidemiologic bubble. As these aging AAS users reach middle age, and enter the age of risk for various medical complications of AAS, it appears likely that we may witness a substantial rise in public health problems associated with illicit AAS use. By analogy, imagine that widespread smoking of cigarettes did not appear until 1980, and that most of the older cigarette smokers were only approaching middle age today. In this hypothetical scenario, one might hear of occasional cases of lung cancer or emphysema, but the full magnitude of risk conferred by cigarette smoking would not yet be appreciated. It is possible that an analogous situation exists with AAS use, and that we will witness a substantial increase in AAS-induced medical problems over the next decade or two.

By contrast with the West, AAS use is rare in most Asian countries such as China or Japan. This difference is likely attributable to differences between Eastern and Western cultural attitudes toward muscularity and masculinity. In the West, muscularity has been a feature of male heroes and gods for thousands of years, as illustrated, for example, by Greek statues of Herakles and the other Titans, or the muscular gods of Norse mythology such as Thor and Vulcan. By contrast, China and Japan have no comparable ancient muscular gods or heroes, and muscularity is not associated with masculinity in the same way. In Confucian tradition, for example, a masculine man is someone who has intellect, integrity, and force of character, but not bulging muscles. Similarly, Japanese gods and heroes are typically shown fully clothed in artwork and statuary, without any of the added muscle that one would see in Western art or statues. Thus, it seems likely that differences in cultural traditions have helped to insulate the East against widespread illicit AAS use.

AAS DEPENDENCE

AAS are usually consumed in courses, colloquially called cycles, which typically extend from a few weeks to a few months in duration, separated by drug-free intervals. It is widely believed by AAS users that the drugs are more effective when taken in discrete cycles than when taken on a continuous basis, perhaps because there is less tolerance to their muscle-building effects, although it is not clear that this belief is scientifically justified. Another important consideration is the fact that exogenous AAS suppress the body's hypothalamic–pituitary–gonadal (HPG) axis, thus suppressing endogenous testosterone production. When AAS are used in cycles, the HPG axis can recover during the spaces between cycles, thus restoring endogenous testosterone production to normal.

Many illicit AAS users are content to ingest only a few cycles of AAS in a lifetime, with a cumulative total lifetime AAS exposure of only 6–12 months. It has become increasingly recognized, however, that AAS can create a dependence syndrome, in which some individuals begin to use AAS almost continuously, with little or no space between cycles, often for years of cumulative lifetime exposure, despite adverse medical, psychiatric, and social consequences. AAS dependence differs in some ways from dependence upon classical drugs of abuse, because AAS do not generally produce a high of intoxication in the manner of drugs such as alcohol, cannabis, or opioids. Thus, AAS do not yield the immediate reward of classically intoxicating drugs, but instead a delayed reward of increased muscularity and enhanced bodily appearance. We will discuss the mechanisms of this dependence syndrome in greater detail in the next section below.

Field studies of illicit AAS users in the United States, United Kingdom, and Australia have attempted to estimate the prevalence of AAS dependence in various populations. On average, these studies estimate that about 30% of male AAS users exhibit a dependence syndrome, meeting criteria for dependence similar to those postulated for other forms of substance dependence in the standard diagnostic nomenclature. However, it should be recognized that this estimate may be influenced by selection bias, since individuals with AAS dependence are perhaps more likely to be encountered in field studies recruiting study participants from gymnasia and other similar venues, whereas nondependent individuals, who perhaps used AAS only briefly, might be

underrepresented in such samples. Nevertheless, even if one assumes that the prevalence of dependence is substantially below 30%, it still appears that well over 1 million individuals, and perhaps even several million worldwide, have developed this syndrome. From a public health perspective, this group is of greatest interest, since it likely accounts for the great majority of the morbidity and possible mortality associated with AAS use in much the same manner as other drugs, such as alcohol, where the majority of the pathology is concentrated in the minority of users who display the highest levels of consumption.

MECHANISMS OF AAS DEPENDENCE

AAS dependence has not been studied in nearly as much detail as other forms of classical drug dependence, and indeed arguably represents the least-studied major form of substance dependence in the world. Nevertheless, available studies suggest that there may perhaps be at least three separate mechanisms whereby AAS dependence can develop. As summarized in Fig. 75.2, these include (1) an anabolic pathway, where individuals with muscle dysmorphia might become dependent on AAS for their muscle-building effects; (2) an androgenic pathway, where men might use AAS to self-treat hypogonadism from AAS withdrawal; and (3) a hedonic pathway, where AAS dependence may arise via mechanisms shared with classical addictive drugs. Each pathway, if supported, would suggest specific clinical treatments (see Fig. 75.2).

The anabolic pathway arises from the ability of AAS to stimulate muscle growth. A substantial literature has shown that individuals with concerns about their body image and muscularity may be particularly vulnerable to using AAS, since these drugs answer to these concerns and provide a rapid and often dramatic increase in muscularity. In particular, it has been shown that a substantial number of men suffer from a form of body dysmorphic disorder termed muscle dysmorphia, in which they become pathologically concerned that they do not look muscular enough. Muscle dysmorphia has also sometimes been called reverse anorexia nervosa, because individuals with this disorder may perceive themselves in the mirror as small and frail, even though they are actually large and muscular, which is the reverse of the pattern seen in classical anorexia nervosa, where girls may perceive themselves as being fat when they are actually thin or even emaciated.

At first it might be assumed that AAS use would help to resolve individuals' concerns about muscularity, but paradoxically, such individuals may often become even more obsessed with muscularity as they grow bigger on AAS, such that they never feel muscular enough. Therefore, they continue to ingest repeated cycles of AAS, sometimes in escalating doses, and may develop anxiety if they begin to lose even a small amount of muscle upon discontinuing AAS. For example, it is not uncommon to hear anecdotally from AAS users that they resumed using AAS much more quickly than intended after a previous cycle, because they developed intolerable anxiety that they might be losing muscular size.

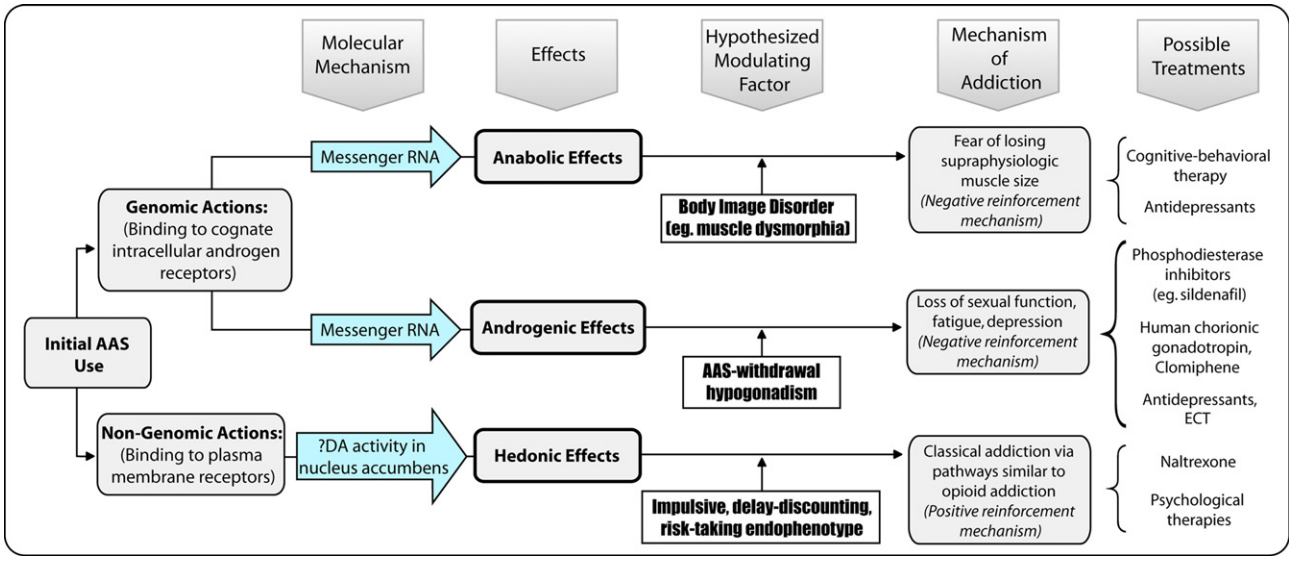


FIGURE 75.2 A model illustrating three possible mechanisms for the evolution of AAS dependence, together with potential treatment strategies to address these mechanisms. DA, dopamine. *Adapted from Kanayama, G., Brower, K.J., Wood, R.I., Hudson, J.I. and Pope, H.G. (2010). Treatment of anabolic-androgenic steroid dependence: emerging evidence and its implications.* Drug and Alcohol Dependence *109, 6–13.*

Although there have been no specific studies addressing the treatment of muscle dysmorphia associated with AAS dependence, many studies have addressed the treatment of other forms of body dysmorphic disorder. It has been found that both cognitive behavioral therapy and serotonergic antidepressants (e.g. serotonin reuptake inhibitors and norepinephrine-serotonin reuptake inhibitors) show efficacy in body dysmorphic disorder, and thus it follows that these modalities may also be helpful for AAS-dependent individuals who experience pathological body image concerns. However, these speculations remain to be tested.

The androgenic pathway to AAS dependence arises from the neuroendocrine effects of AAS. As mentioned earlier, testosterone and other AAS suppress the HPG axis, leading to testicular atrophy and decreased production of testosterone and spermatozoa in males. If a man stops AAS, particularly if he stops AAS abruptly after a long cycle, he will be very likely to become hypogonadal, because the HPG axis may require weeks or months to regain normal functioning. Indeed, several reports have described AAS users who exhibited hypogonadism lasting for more than a year after discontinuing AAS.

Although AAS users rarely describe an explicit craving for AAS in the manner of individuals experiencing, say, alcohol or cocaine withdrawal, they frequently experience dysphoric symptoms associated with AAS-withdrawal hypogonadism. These may include fatigue, decreased sex drive, hypersomnia, loss of appetite, decreased enthusiasm for activities, and in some cases severe depression. It appears that serious depressive symptoms are an idiosyncratic effect of AAS-withdrawal hypogonadism, in that a majority of individuals do not develop marked depression, even when profoundly hypogonadal, whereas a small minority develop pronounced depression, sometimes associated with suicidal ideation or outright suicide attempts. In any case, the dysphoric symptoms associated with hypogonadism, whether they be mild or severe, likely induce some individuals to resume AAS use in order to self-treat these symptoms, thus leading to repeated AAS use and ultimately to dependence.

Because illicit AAS use is a relatively new phenomenon, as noted above, many physicians are unfamiliar with the treatment of hypogonadism in AAS users. An evolving literature suggests that it is important to recognize and promptly treat hypogonadism in these individuals, especially when they have expressed a desire to stop their drug use, because such treatment may help to prevent them from resuming AAS. A variety of neuroendocrine treatments exist for this purpose. For example, exogenously administered human chorionic gonadotropin mimics the effects of pituitary luteinizing and follicle-stimulating hormones, thus stimulating the testis to resume production of testosterone and spermatozoa more rapidly. The drug clomiphene in turn mimics the effects of hypothalamic-releasing hormones and thus stimulates pituitary secretion of luteinizing and follicle-stimulating hormones, also aiding in the restoration of normal HPG axis functioning. Although many psychiatrists and other substance abuse professionals are unfamiliar with such neuroendocrine treatments, patients with AAS-withdrawal hypogonadism can easily be referred to endocrinologists for management of this problem.

A third pathway to AAS dependence might be called the hedonic pathway. Persuasive evidence for this pathway arises from animal studies, which have found that laboratory animals show a preference for AAS, even though these animals are presumably not concerned with their muscularity or with the threat of becoming hypogonadal. For example, rats and mice display conditioned place preference for AAS, in that they prefer to spend time in an environment where they have previously received AAS as compared to a comparable environment where they have not. Even more compelling are recent studies showing that male hamsters self-administer testosterone and other AAS, even to the point of death in some cases. Importantly, self-administration occurs even when testosterone is administered directly into the cerebral ventricles; this finding indicates that self-administration is not dependent on some peripheral effect of testosterone (e.g. via an effect on muscle or joint pain), but is attributable to a central effect in the brain.

The mechanism of AAS self-administration in animals remains incompletely understood. However, several lines of evidence suggest that this self-administration is mediated by binding of drug to plasma membrane receptors in the brain, rather than by the classic genomic effects of AAS whereby AAS enter the cell nucleus and bind to intracellular androgen receptors. More specifically, animal studies suggest that there may be metabolites of the parent AAS that bind to receptor sites on the nucleus accumbens, producing hedonic effects analogous to those produced by classical addictive drugs. Interestingly, hamsters intoxicated with testosterone after self-administration exhibit a state somewhat similar to opioid intoxication, suggesting possibly similar mechanisms. Furthermore, pretreatment of hamsters with opioid antagonists, such as naltrexone, blocks intracerebroventricular testosterone self-administration, again suggesting that opioidergic mechanisms may mediate this process.

Several observations in human AAS users appear to parallel the observations from animal studies. In one of the earliest case reports of AAS dependence, the investigators found that an AAS-dependent individual exhibited opioid-like withdrawal symptoms when

challenged with naloxone, even though this man had apparently never used opioids. In another more recent study, the investigators evaluated consecutive men admitted to an opioid detoxification facility and found that nearly 10% of them appeared to have been introduced to opioids via their initial use of AAS. These men typically reported that they had first learned about using opioids from friends at the gym, and had purchased their first illicit opioids from the same drug supplier from whom they had purchased AAS. Another study examined sequential men admitted to a generic substance dependence treatment facility and asked all of them about prior history of AAS use. Among patients reporting opioids as their drug of choice, a past history of AAS use was much more common than among patients reporting other forms of drug use such as alcohol dependence. Finally, a recent field study compared AAS-dependent men, nondependent AAS users, and comparison weightlifters reporting no history of AAS use. The men with AAS dependence reported a strikingly higher lifetime prevalence of opioid abuse or dependence than the other groups, again suggesting that AAS and opioids possibly produced similar hedonic effects in these men. Interestingly, the chronological sequence of AAS and opioid use in these men was variable; opioid abuse or dependence occurred before, during, and after AAS use in different individuals, suggesting that neither form of drug abuse necessarily led to the other, but rather that they might arise from a common underlying mechanism (Fig. 75.3).

The implications of these latter findings for treatment of AAS dependence remain uncertain, because there are as yet no major studies of treatment of AAS dependence in the literature. Certainly, it seems reasonable to speculate that therapies effective for opioid and other substance dependence might also be effective for AAS dependence. Such therapies might include motivational therapies in the early stages of treatment to address denial and minimization, followed by psychological therapies such as supportive-expressive therapy, contingency management, and behavioral couples therapy or behavioral family counseling. It seems likely that couples therapy might be particularly useful in such cases, since women may suffer physical and emotional abuse from male AAS-using partners.

Another, more provocative possibility is that use of an opiate antagonist, such as naltrexone, might be helpful for AAS dependence. It has already been shown that long-acting preparations of naltrexone may be effective in the treatment of alcohol dependence, perhaps by blocking the effects of endogenous opioids that are released following the ingestion of alcohol. By analogy, naltrexone might also block endogenous opioid-like effects mediated by AAS. To our knowledge, however, this approach has not yet been attempted in individuals with AAS dependence.

HAZARDS OF AAS DEPENDENCE

Although many forms of addiction are well established as harmful and deserving of treatment, the hazards of AAS dependence are still poorly understood. Indeed, widespread opinion in the AAS-using community and in websites and discussion groups from the bodybuilding underground has questioned whether AAS use is seriously harmful. Similarly, prominent reports in the lay media and even articles from some scientists have argued that warnings about the dangers of AAS are exaggerated; some have sometimes even labeled these warnings as steroid hysteria. Indeed, AAS use is sometimes portrayed as part of a healthy athletic lifestyle, rather than as simply a form of illicit drug

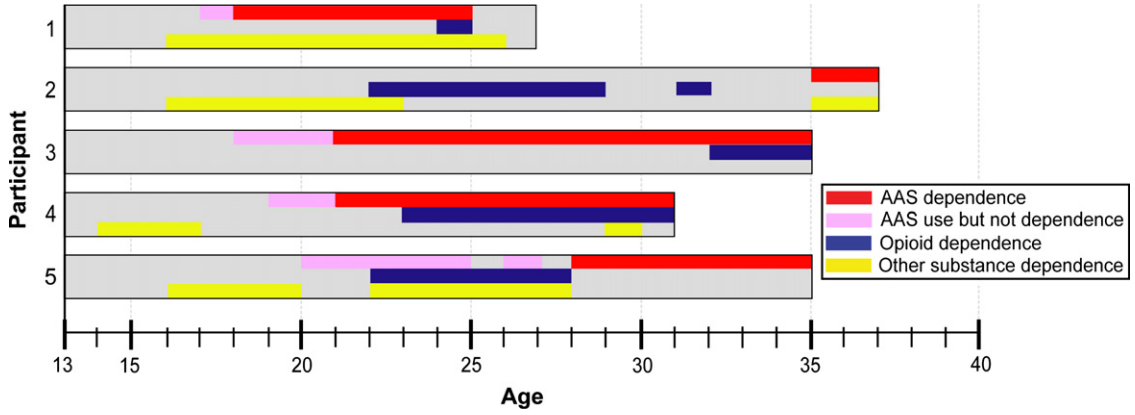


FIGURE 75.3 Illustrations of the time course of AAS abuse and dependence, together with other forms of substance abuse or dependence, in five representative men with AAS dependence. *Adapted from Kanayama, G., Hudson, J.I. and Pope, H.G. Jr. (2009). Features of men with anabolic-androgenic steroid dependence: a comparison with nondependent AAS users and with AAS nonusers.* Drug and Alcohol Dependence *102, 130–137.*

use – a portrayal that belies the fact that many AAS users ingest large amounts of classical illicit drugs as well.

However, these lenient attitudes about AAS have been increasingly challenged by evidence of serious medical and psychiatric adverse effects from long-term exposure to these drugs. Perhaps the most extensive medical evidence surrounds the cardiovascular effects of AAS. There is also evidence of adverse medical effects on other organ systems, such as the liver and kidneys. In addition there is substantial evidence that AAS can produce adverse psychiatric effects.

Cardiovascular Effects

AAS produce a range of cardiovascular effects, which appear to be responsible for substantial morbidity and even premature mortality. First, AAS can produce hypertension in some individuals, especially when these drugs are used in massively supraphysiologic doses, as is often the case when they are used illicitly. AAS also can increase red blood cell production, resulting in polycythemia, accompanied by an increased risk of thrombosis. In both the scientific literature and the anecdotal experience of the authors, these effects may lead to apparently thrombotic events, such as transient ischemic attacks, cerebrovascular accidents, and myocardial infarction.

AAS may also affect cardiac electrical conduction, although the literature is inconsistent with regard to this finding. In one recent study, for example, AAS users showed significantly longer PR intervals and QRS intervals as compared to otherwise similar weightlifters who denied AAS use. This finding suggests that AAS users may display impaired cardiac electrical propagation. Given the association between such electrocardiographic abnormalities and the risk for pathologic arrhythmias in other populations, together with reports of unexplained arrhythmias and sudden death in AAS users, this finding is of some concern.

Even more ominous are findings associating AAS use with myocardial toxicity. A growing literature has suggested that AAS use may be associated with impaired myocardial contraction, as reflected, for example, in a decreased left ventricular ejection fraction. One recent study found that 10 (83%) of 12 long-term AAS users displayed a left ventricular ejection fraction below the accepted limit of normal (55%) as compared to only one (14%) of seven comparison weightlifters reporting no AAS usage. AAS users have also been reported to exhibit diastolic impairment, as reflected by lower early diastolic transmitral blood flow velocity (E-wave) and early peak tissue velocity (E') with preserved or increased late diastolic filling, as evidenced by late diastolic transmitral blood flow velocity (A-wave) and peak tissue velocity (A'). The clinical significance of these findings remains incompletely understood, but a number of anecdotal reports in the literature have described frank cardiomyopathy in AAS users, and it seems plausible that some of this pathology may be irreversible.

Finally, AAS use is associated with dyslipidemia. In individuals ingesting markedly supraphysiologic doses of AAS, especially orally active AAS, levels of high-density lipoprotein cholesterol may fall dramatically – sometimes to less than 10 mg dL^{-1} – whereas levels of low-density lipoprotein cholesterol may rise. These alterations in lipid levels almost certainly have serious atherogenic potential, although again the full clinical significance of these effects remains unknown. Certainly there are many anecdotal reports of myocardial and cerebral infarctions in individuals with known or probable AAS use who were only in their 30s or 40s, and it seems likely that some of these may have been attributable to premature atherosclerotic changes. One small preliminary imaging study has shown striking increases in coronary artery calcium in bodybuilders with long-term AAS exposure, but these findings remains to be replicated in larger investigations.

It is difficult to estimate the overall threat posed by the many cardiovascular effects of AAS use. One study in Finland examined 62 men who had placed highly in power lifting contests in 1977–1982 and compared them to matched population controls. Eight of the 62 men had died – a rate much greater than in the general population – and three of these died of myocardial infarction, which the authors speculated was likely associated with AAS exposure. Similarly, a Swedish pathological study found chronic cardiac changes in 12 (35%) of 34 medicolegally investigated deaths of male users of AAS.

As discussed earlier in this chapter, the older members of the illicit AAS-using population are only now entering middle age, and thus it seems likely that the above cardiac effects will become more manifest as this population moves into the age of risk. Thus, the cardiac effects of long-term AAS exposure represent an area urgently in need of further study.

Effects on Other Organ Systems

Long-term exposure to supraphysiologic doses of AAS may also produce deleterious effects on other organ systems. For example, a substantial literature has shown that orally active AAS may exhibit hepatotoxicity, as evidenced by inflammatory changes, peliosis hepatis (the formation of blood-filled cysts in the liver), and the development of various forms of liver cancer. Although the association between these effects and AAS use is very well documented, these effects appear to be relatively rare. Interestingly, public warnings about

the dangers of AAS often emphasize hepatotoxicity, out of proportion to the actual prevalence of these phenomena. Our experience suggests that a possible reason for this misplaced emphasis is the fact that many AAS users exhibit substantial elevations of liver enzymes, such as alanine aminotransferase, aspartate aminotransferase, and lactic dehydrogenase. However, these same enzymes are also present in skeletal muscle, and muscle trauma from a hard workout in the gymnasium can produce marked elevations in these indices, even in individuals with no liver disease. Thus, it seems likely that many practitioners have misinterpreted elevations in these enzymes as suggestive of AAS-induced hepatotoxicity when in fact they are not. A more specific test is the γ-glutamyl transaminase (GGT), an enzyme present exclusively in liver and not in skeletal muscle. In studies of AAS users conducted by the authors, GGT has almost always been within the normal range, suggesting that AAS-induced hepatic inflammation is rare.

A potentially more serious concern is nephrotoxicity. Massive muscle breakdown occurring in AAS-using weightlifters may lead to rhabdomyolysis and myoglobin-induced renal failure. One recent report has described focal segmental glomuleronephritis in ten long-term AAS users, possibly attributable to a combination of postadaptive glomerular changes, driven by increased muscle mass, and to a potential direct nephrotoxic effect of AAS. These renal complications may be largely reversible if AAS are discontinued, but may go unrecognized in AAS users for prolonged periods.

The effects of AAS on the prostate are less clear. Although it has been widely believed that supraphysiologic levels of testosterone or other AAS might potentiate or exacerbate prostate cancer, there are as yet only two individual case reports of prostate cancer in middle-aged bodybuilders with possible AAS use, and both of these reports appeared about 20 years ago. The absence of more recent reports argues against an association between AAS use and prostate cancer. Moreover, some authorities have recently argued that it is a myth that androgen administration can stimulate prostate cancer, even though it is universally acknowledged that reduction of androgen levels can slow the progression of existing prostate cancer. In any event, at present, it appears that the dangers from the prostatic effects of long-term AAS exposure rank well below the dangers from the cardiac and renal effects of these drugs.

Psychiatric Effects

A substantial literature has now described psychiatric effects of AAS. Unlike the medical effects discussed above, many psychiatric effects are not contingent upon long-term exposure and may occur even in some individuals with relatively brief AAS experience. The principal psychiatric effects of AAS include manic or hypomanic syndromes during AAS exposure, and depression during AAS withdrawal.

Manic or Hypomanic Effects

As early as 1985, it was reported that AAS use could apparently induce hypomanic symptoms, including euphoria, hyperactivity, hypersexuality, and impaired judgment. Many studies appeared over the next 10 years, documenting hypomanic or frankly manic symptoms in some AAS users, occasionally even accompanied by psychotic symptoms, such as grandiose or paranoid delusions. These effects appear to be dose related; individuals using only modestly supraphysiologic doses of AAS (up to 300 mg of testosterone or equivalent per week) have rarely been reported to show such symptoms, while individuals in the mid-range of dosage (300–1000 mg of testosterone equivalent per week) show a somewhat higher prevalence, and those in the upper range (greater than 1000 mg of testosterone equivalent per week) appear most likely to exhibit such effects. In one study, for example, 7 (28%) of 25 illicit AAS users taking more than 1000 mg of testosterone per week reported symptoms meeting criteria for a hypomanic or manic episode, as opposed to only 1 (7%) of 14 individuals using less than 300 mg per week.

Starting in about 1990, a number of reports also appeared suggesting that hypomanic symptoms in AAS users might include unusual irritability, aggression, or even violence. Several reports described individuals with no apparent past history of psychiatric disorder, violent behavior, or criminality who became uncharacteristically aggressive and committed violent crimes, including murder or attempted murder, while under the influence of AAS. In several forensic cases around the United States, the issue of involuntary intoxication with AAS was even raised as a possible defense, albeit with mixed results.

Several writers have questioned whether AAS play an etiologic role in the development of hypomanic or aggressive symptoms. For example, it might be argued that aggressive individuals are more prone to take AAS in the first place, and that their behavior does not represent a biological effect of the drugs. Similarly, expectational factors or sociocultural effects of being in the gym subculture might plausibly explain such behavior. However, these concerns have now been addressed by four blinded investigations, in which supraphysiologic doses of AAS (equivalent to 500 mg of testosterone per week or greater) were administered to normal volunteers. In the four studies, collectively comprising 105 individuals administered AAS, 5 (4.8%) developed hypomanic or manic symptoms

during AAS exposure, even though they did not exhibit such symptoms on placebo or at other times in their lives. Since these observations were obtained under blinded conditions, it appears that these effects cannot be explained purely on the basis of psychological or expectational factors, and indeed represent a biological effect of AAS. However, it remains unclear why occasional individuals are highly vulnerable to these mood-altering effects, while most individuals do not develop such symptoms despite ingesting identical doses under similar conditions.

Major Depression

As we have noted earlier, major depression may sometimes develop in the context of AAS-withdrawal hypogonadism, and this effect, like the hypomanic effects just discussed, appears to be an idiosyncratic phenomenon that affects only a minority of individuals. Again, the reasons for this idiosyncratic response remain unknown, but it is well recognized that some men will occasionally develop pronounced depressive symptoms, sometimes associated with suicidal ideation or occasionally even completed suicide, typically after stopping a prolonged cycle of AAS use. One report has described successful treatment of four cases of AAS-withdrawal depression using the antidepressant fluoxetine; another report described a case resistant to antidepressant medications, but ultimately responsive to electroconvulsive therapy. As we have noted earlier, it is probably important to treat AAS-withdrawal hypogonadism with neuroendocrine therapy, while remaining alert to the possibility that some individuals will occasionally exhibit serious depression requiring psychiatric intervention over and above endocrine treatment.

FUTURE PROSPECTS

Of the many forms of addiction discussed in this volume, dependence on illicit AAS is one of the least studied and least understood. This dearth of evidence is partly attributable to the fact that AAS use is a relatively recent form of illicit drug use, as discussed above, and also attributable to the fact that AAS users rarely seek treatment. Indeed, AAS users are frequently distrustful of health professionals and often report that they have never disclosed their AAS use to any doctor that they have seen. Similarly, there have been relatively few investigators who have successfully recruited and personally evaluated large numbers of illicit AAS users. Although there are numerous animal studies of AAS, anonymous questionnaire studies asking about AAS use, Internet surveys of AAS users, and studies on programs to prevent AAS use, there are only a modest number of studies in which actual illicit users have been assessed in person.

However, it is possible that this situation may change over the next decade, as increasing numbers of men with AAS dependence grow old enough to experience adverse effects from prolonged exposure to these drugs, and consequently may be induced to seek treatment when they would not have done so earlier. Such men might perhaps come to the attention of cardiologists as a result of cardiomyopathy or atherosclerotic disease, to nephrologists as a result of real disease, or to endocrinologists and psychiatrists as a result of prolonged hypogonadism and possible associated major depressive disorder. It will likely become important for clinicians in these various disciplines to become familiar with AAS dependence and its effects, so that they can quickly identify such patients and refer them for appropriate treatment.

At present, as we have noted above, experience with treatment of AAS dependence is very limited, although possible modalities of treatment are already apparent, as also discussed earlier. The next decade will likely see a substantial increase in our knowledge of this area, as this form of dependence becomes better known.

SEE ALSO

Athletes and Substance Use

List of Abbreviations

AAS	anabolic-androgenic steroids
AIDS	acquired immunodeficiency syndrome
GGT	γ-glutamyl transaminase
HIV	human immunodeficiency virus
HPG axis	hypothalamic–pituitary–gonadal axis

Further Reading

Baggish, A.L., Weiner, R.B., Kanayama, G., Hudson, J.I., Picard, M.H., Hutter Jr., A.M., Pope Jr., H.G., 2010. Long term anabolic androgenic steroid use is associated with left ventricular dysfunction. Circulation: Heart Failure 3, 472–476.

Brower, K.J., 2009. Anabolic steroid abuse and dependence in clinical practice. The Physician and Sportsmedicine 37, 1–11.

Gruber, A.J., Pope Jr., H.G., 2000. Psychiatric and medical effects of anabolic-androgenic steroid use in women. Psychotherapy and Psychosomatics 69, 19–26.

Kanayama, G., Boynes, M., Hudson, J.I., Field, A.E., Pope Jr., H.G., 2007. Anabolic steroid abuse among teenage girls: an illusory problem? Drug and Alcohol Dependence 88, 156–162.

Kanayama, G., Brower, K.J., Wood, R.I., Hudson, J.I., Pope Jr., H.G., 2010. Treatment of anabolic-androgenic steroid dependence: emerging evidence and its implications. Drug and Alcohol Dependence 109, 6–13.

Kanayama, G., Hudson, J.I., Pope Jr., H.G., 2008. Long-term psychiatric and medical consequences of anabolic-androgenic steroid

abuse: a looming public health concern? Drug and Alcohol Dependence 98, 1–12.

Parssinen, M., Seppala, T., 2002. Steroid use and long-term health risks in former athletes. Sports Medicine 32, 83–94.

Pope Jr., H.G., Katz, D.L., 2003. Psychiatric effects of exogenous anabolic-androgenic steroids. In: Wolkowitz, O.M., Rothschild, A.J. (Eds.), Psychoneuroendocrinology: The Scientific Basis of Clinical Practice. American Psychiatric Press, Washington, DC, pp. 331–358.

Tan, R.S., Scally, M.C., 2009. Anabolic steroid-induced hypogonadism – towards a unified hypothesis of anabolic steroid action. Medical Hypotheses 72, 723–728.

Thiblin, I., Petersson, A., 2005. Pharmacoepidemiology of anabolic androgenic steroids: a review. Fundamentals of Clinical Pharmacology 19, 27–44.

Wood, R.I., 2008. Anabolic-androgenic steroid dependence? Insights from animals and humans. Frontiers in Neuroendocrinology 29, 490–506.

CHAPTER

76

Prescription and Over-the-Counter Medications

Jessica L. Eaddy

Clinical Neuroscience Division, Medical University of South Carolina, Charleston, SC, USA

OUTLINE

INTRODUCTION

The nonmedical use of prescription and over-the-counter (OTC) medications is a serious public health concern. Prescription medications are drugs approved by the US Food and Drug Administration (FDA) that require an individualized prescription written by a licensed health care professional. The most commonly abused prescription medications, which are the focus of this chapter, include pain relievers (i.e. opioids), stimulants, and sedatives/tranquilizers. OTC medications refer to all FDA-approved medications available directly to consumers without a prescription. These drugs are considered to be relatively safe, effective, and involve minimal risk when taken as directed. OTC medications containing pseudoephedrine, antihistamines, or dextromethorphan (DXM), in particular, are discussed. See Table 76.1 for a list of generic and trade names of common prescription and OTC medications.

Prescription medications are some of the most commonly abused substances. Each year, the National Survey on Drug Use and Health (NSDUH) assesses persons in the United States aged 12 years or older regarding the use of illicit drugs. The NSDUH asks about marijuana/hashish, cocaine (including crack), inhalants (e.g. nitrous oxide, amyl nitrite, cleaning fluids, gasoline, spray paint, other aerosol sprays, glue), hallucinogens (e.g. lysergic acid diethylamide (LSD), phencyclidine (PCP), peyote, mescaline, psilocybin mushrooms, ecstasy), heroin and prescription drugs (e.g. pain relievers, tranquilizers, stimulants, sedatives). In 2009, approximately 22 million Americans (i.e. 8.7% of the United States population) reported using an illicit drug in the past month. With regard to the nonmedical use of prescription medications, 7.0 million Americans (e.g. 2.8% of the United States population) reported past month use. Some data

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00076-0

TABLE 76.1 Generic and Trade Names of Common Prescription and OTC Medications

Generic name	Brand names
Opioids	
Buprenorphine	Subutex, Buprenex
Codeine	Phenergan with Codeine, Empirin with Codeine, Robitussin A-C, Tylenol with Codeine, Actifed with Codeine, C Tussin, Cheratussin, Zodryl
Hydrocodone	Vicodin, Vicoprofen, Anexsia, Dolorex Forte, Hycet, Liquicet, Lorcet, Lortab, Maxidone, Norco, Polygesic, Xodol, Zydone, Hycodan
Hydromorphone	Dilaudid, Palladone, Hydrostat, Hydromorph
Fentanyl	Ionsys, Duragesic, Sublimaze, Actiq, Fentora
Levacetylmethadol (LAAM)*	OrLAAM*
Levorphanol	Levo-Dromoran
Meperidine	Demerol
Methadone	Diskets, Dolophine, Methadose
Nalbuphine	Nubaine
Opium tincture	Generic only; no brands available
Oxycodone	Dazidox, ETH-Oxydose, Endocodone, OxyIR, Oxycontin, Oxyfast, Percolone, Roxicodone
Pentazocine	Fortral, Talwin, Talacen, Fortwin
Propoxyphene	Darvon, Darvocet, PP-Cap
Remifentanil	Ultiva
Sufentanil	Sufenta
Tapentadol	Nucynta
Tramadol	Ultram, Ultracet
Sedatives-hypnotics	
Barbiturates	
Amobarbital	Amytal
Glutethimide	Doriden
Mephobarbital	Mebaral
Methaqualone*	Quaaludes
Pentobarbital Sodium	Nembutal
Phenobarbitol	Solfoton, Luminal
Secobarbital	Seconal
Thiopental	Pentothal, Trapanal
Benzodiazepines	
Alprazolam	Xanax, Niravam
Chlorazepate	Gen-Xene, Tranxene
Chlordiazepoxide	Librium
Clonazepam	Klonopin
Diazepam	Valium, Valrelease, Dizac, Zetran, Diastat
Estazolam	ProSom
Flunitrazepam*	Rohypnol*
Flurazepam	Dalmane
Halazepam	Paxipam
Lorazepam	Ativan
Midazolam	Versed, Dormicum, Hypnovel
Oxazepam	Serax
Prazepam	Centrax
Quazepam	Doral
Temazepam	Restoril
Triazolam	Halcion
Nonbarbiturate sedative-hypnotics	
Chloral hydrate	Notec, Somnote, Felsules
Eszopiclone	Lunesta
Ramelteon	Rozerem
Zaleplon	Sonata
Zolpidem	Ambien
γ-Hydroxbutyrate (GHB)*	Xyrem*
Stimulants	
Dexmethylphenidate	Focalin
Methylphenidate	Ritalin, Concerta, Metadate, Methylin, Daytrana
Amphetamines	
Amphetamine–dextroamphetamine mixtures	Adderall, Biphetamine* Dexamphetamine
Dextroamphetamine sulfate	Dexedrine, Dextrostat, ProCentra
Lisdexamfetamine dimesylate	Vyvanse

TABLE 76.1 Generic and Trade Names of Common Prescription and OTC Medications—cont'd

Generic name	Brand names
Methamphetamine hydrochloride	Desoxyn
Analeptics	
Doxapram	Dopram
Modafinil/armodafinil	Provigil, Nuvigil
Anorexiants	
Benzphetamine	Didrex
Diethylpropion	Tenuate
Fenfluramine*	Redux*
Phendimetrazine tartrate	Bontril
Phentermine*	Adipex,* Fastin*
Sibutramine*	Meridia,* Reductil*
OTC drugs	
Decongestants	
Ephedrine**	ePHEDrine Sulfate, Am-Ephed, Amesec, Amphetrazine, Kondon's Nasal, Mudrane, Primatene, Theodrine, and others
Phenylephrine	Sudafed PE, Sudogest PE, Dimetapp, Triaminic, Vick's Sinex, and others
Pseudoephedrine**	Sudafed, Sudogest, Dimetapp (Maximum Strength), Sudophed, Simply Stuffy, and others
Antihistamines	
Dimenhydrinate	Dramamine, Tryptone
Diphenhydramine	Benadryl, Simply Sleep, Nytol, Midol, AllerMax, Theraflu strips, and others
Doxylamine	Aldex, Doxytex, Nitetime, Nytol, Sleep Aid (Doxylamine), Unisom
Diet aids	
Ephedra*	Diet Pep,* Metabolife, Gen-Fen, Herbal Majic, Twinlab Ripped, Trim Easy Plus
Antitussives	
Dextromethorphan	Coricidin, Robitussin, Delsym, Alka-Seltzer, Dimetapp, Triaminic, Theraflu, Nyquil

** The drug has been removed from the market.*
*** Regulation varies by state; minimum federal requirement that consumer present photo identification to purchase.*

suggest that among youth aged 12–17 years old, females are more likely than males to use prescription medications nonmedically, but among individuals 18 years or older, males report higher prescription medication use than females. This gender difference may also vary by the type of medication.

Each day in the United States, an estimated 8500 individuals initiate illicit drug use for the first time (see Fig. 76.1). Over half of all new initiates are less than 18 years old and approximately half are female. Among new initiates, a substantial amount (28.6%) begin illicit drug use by consuming a prescription medication, the most common of which are pain relievers (17.1%).

In contrast to other commonly abused substances in society, such as alcohol and cocaine, prescription and OTC medications are unique in that their consumption is endorsed or prescribed by health care professionals. For many individuals, these medications serve as a primary treatment for certain medical conditions and represent a legitimate means of improving physical health and well-being. However, this is only true when the medications are used exactly as directed. Because these medications are FDA-approved and many are prescribed by health care professionals, some individuals develop a false sense of safety regarding their use and erroneously believe that the severity and risk of negative side effects is lower with prescription and OTC medications compared with other substances. Teenagers and young adults, who often have a sense of being invulnerable, are particularly susceptible to this erroneous perception of prescription and OTC drugs as being safe.

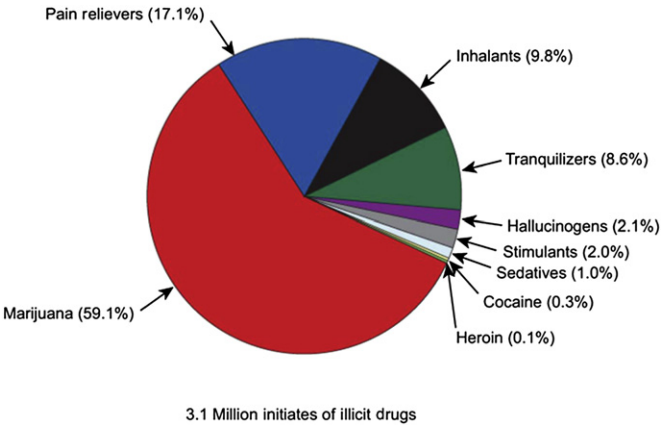


FIGURE 76.1 Specific drugs associated with initiation of illicit drug use among 3.1 million past year illicit drug initiates aged 12 years or older. From Substance Abuse and Mental Health Services Administration (2010). Results from the 2009 National Survey on Drug Use and Health: Volume I. Summary of national findings. Rockville, MD: Office of Applied Studies, NSDUH Series H-38A, HHS Publication No. SMA 10-4856 Findings.

RISK AND PROTECTIVE FACTORS

Table 76.2 includes a list of factors that may increase the risk of nonmedical prescription and OTC medication use. For example, individuals who initiate substance use at an early age, have a family history of alcohol or drug addiction, and aged 18–25 years are at increased risk of misusing prescription and OTC medications. Individuals with other substance use disorders or co-occurring psychiatric symptoms such as depression, anxiety, attention deficit hyperactivity disorder (ADHD) and eating disorders are also at increased risk. Such individuals may use the medications in an attempt to self-medicate their psychiatric symptoms. Women, in particular, may use stimulants in an attempt to lose or maintain weight. A history of childhood trauma (e.g. sexual or physical assault) or neglect may also increase the risk of nonmedical prescription and OTC medication use. Numerous studies demonstrate a strong correlation between exposure to early-life adversity, such as trauma and violence, and elevated risk for addiction as an adult. Finally, living with someone who uses prescription and OTC medications may also increase risk. Data regarding pain medication misuse show that one of the most common sources for obtaining pain relievers is family and friends. Young adults may be given a prescription medication by well-intentioned but ill-informed family members to help with a physical condition (e.g. a mother may give her teenage daughter one of her own prescription pain relievers to sooth a migraine headache or menstrual cramps). Furthermore, risk is increased when living with a user of prescription medications, even when that use is legitimate. Simply having access to these drugs is a risk factor. Caregivers should store all prescription and OTC medications securely to keep others in the home safe. Appropriate storage involves keeping medications somewhere out of reach (e.g. in a high, cool, and dry location) rather than in a medicine cabinet where anyone who enters the home can access the drugs.

TABLE 76.2 Risk and Protective Factors

Risk factors
Younger age (e.g. 18–25 years old)
Early initiation of substance use
Family history of addiction
Easy access to drugs
Male gender (except among youth, girls outnumber boys in nonmedical use of prescription drugs)
Juvenile delinquency/gang involvement
Societal discrimination (on basis of race, ethnicity, education, sexual orientation, etc.)
Comorbid psychiatric conditions, such as other substance use disorders, mood and anxiety disorders, posttraumatic stress disorder, ADHD, and eating disorders
Low socioeconomic status and poverty
Conflict in interpersonal relationships
History of interpersonal violence, childhood sexual abuse, or other traumas
Sensation-seeking personality characteristics and poor impulse control
Chronic medical problems or frequent surgical procedures (specific to prescription opioids)
Protective factors
Stable family and interpersonal relationships
Strong social support network consisting of nonusing peers
Coping skills (anxiety management, problem solving, emotion regulation skills, etc.)
Religious and spiritual practices
Commitment to academic achievement
Participation in child, adolescent, and family support programs offering tutoring, social competency skills, and alternative positive activities for at-risk youth
Supportive, drug-free friends, and romantic partner

SPECIFIC MEDICATIONS

The following sections briefly address the approved medical uses, mechanisms of action, symptoms of intoxication/withdrawal, and treatment of the most commonly abused prescription medications: opioids, stimulants, and sedative/tranquilizers. Commonly abused OTC medications are briefly discussed at the end.

Prescription Opioids

Medical Uses

Prescription opioids are most often prescribed as analgesics for acute pain relief and degenerative chronic pain conditions (e.g. cancer, rheumatoid arthritis). In addition to their analgesic properties, opioids are also used as anesthetics during surgical procedures and as antitussives to suppress coughing. Various opioids (e.g. methadone, buprenorphine, naltrexone) have also been approved for detoxification and treatment of opioid dependence, as discussed later in this chapter. Although less frequent, prescription

opioids are approved for the treatment of severe diarrhea.

Neurobiology

Opioids work through a complex set of interactions involving multiple neurotransmitter systems within several brain circuits. Chronic use of opioids results in neuroadaptations within these brain areas. Like most drugs of abuse, opioids activate dopaminergic neurons in the mesocorticolimbic and nigrostriatal pathways of the brain. Thus far, studies have focused primarily on the action of opioids within the mesocorticolimbic pathway, which contains dopaminergic neurons originating in the ventral tegmental area (VTA) and projecting into the nucleus accumbens (NAc), amygdala, prefrontral cortex (PFC), and ventral pallidum. The pedunculopontine nucleus is a structure unique to the reinforcing effects of opioid use. Neurotransmitters that specifically activate opioid receptors throughout the body are also important in the acquisition and maintenance of opioid dependence. For example, growing evidence indicates that cannabinoid receptors (CB1s) may play an important role in opioid addiction. One study found that when CB1s were deactivated, rodents no longer self-administered morphine.

Pharmacology

Opioid drugs bind to receptors in the brain that are activated by naturally occurring, endogenous chemicals called opioid peptides. Opioid peptides include (1) endorphins, (2) endomorphins, (3) enkephalins, and (4) dynorphins. Endorphins, specifically β-endorphins originating in the hypothalamus, are responsible for the high or euphoric state produced by opioids. They also produce symptoms of drowsiness, respiratory depression, and analgesia. Enkephalins are spread throughout the body. Within the brain, enkephalins are involved in pain management, respiratory regulation, and may also play a role in the reinforcing properties of opioids. Within the gastrointestinal tract, enkephalins are responsible for digestive function and are associated with the constipating side effect of opioid use. Although dynorphins produce analgesia, they also cause anxiolytic symptoms, anhedonia, confusion, and sweating.

Opioid peptide neurotransmitters can have a variety of psychological and physical effects, depending on what type of receptor(s) they bind to, binding affinity for the receptor, and how strongly they activate the receptor once attached. The exact role of each receptor is still under investigation; however, several generalizations can be made at this point. Mu (μ) opioid receptors are primarily responsible for the subjective euphoria and rewarding properties that users may experience from taking opioid drugs. Blocking or removing μ opioid receptors in rodents dependent on opioids decreases self-administration of opioid drugs and reduces the amount of time rodents spend in areas that remind them of past opioid use (i.e. conditioned place preference). In addition to rewarding effects, μ opioid receptors produce analgesia, pupil constriction, and respiratory depression. Kappa (κ) opioid receptors produce a subjective dysphoric state, pupil constriction, spinal anesthesia, sedation, and stress-induced analgesia. These receptors are also responsible for the hallucinogenic properties of some opioids (e.g. pentazocine, nalbuphine). Dynorphins appear to have a high binding affinity to the κ receptors. The delta (δ) receptors are currently the least understood receptor class, but appear to have analgesic properties.

When an opioid drug binds to a receptor, one of three events can occur. The drug can serve as an antagonist, full agonist, or partial agonist. Drugs that serve as antagonists bind to receptors but do not activate them. Instead, they block other opioid peptides from binding to and activating the receptor. Opioid antagonists that have a high binding affinity for a particular receptor can actually displace opioid agonists from the receptor site, thereby functionally deactivating the receptor and reversing any previous effects. Naltrexone, an opioid antagonist approved for the treatment of opioid dependence, functions in this manner. Drugs that act as opioid full agonists mimic opioid peptides by fully activating the receptors to which they bind. Partial agonists bind to the receptor but do not fully activate it. Pharmacotherapies such as buprenorphine work by competing with agonists at the receptor site and binding tightly to the receptor. However, once attached to the receptor, these partial agonists fail to activate the receptor with the same intensity as an opioid agonist. Partial agonists activate receptors enough to prevent severe withdrawals and cravings, but do not provide the full range of effects that opioids provide, including the rewarding euphoric effects.

Dopamine (DA) is a neurotransmitter critical to modulating the motivational aspects of reward-based behaviors. DA levels are increased in the VTA and NAc when opioid agonists bind to opiate receptors. Opioid agonists trigger a decreased activation response from γ-aminobutyric acid (GABA) interneurons. As these interneurons pump out less GABA (an inhibitory neurotransmitter), the presynaptic dopaminergic neuron is disinhibited. Decreased inhibition allows this dopaminergic neuron to release more DA into the synapse. Increased releases of DA and endorphins are responsible for most of the symptoms of opioid intoxication. Likewise, decreased DA levels within the mesolimbic pathway during opioid

withdrawal are responsible for the dysphoric symptoms following cessation of chronic opioid use. For a detailed discussion of the neuropharmacology of opioids.

Symptoms of Intoxication and Withdrawal

Symptoms of opioid intoxication vary by drug, dosage, and route of administration. Typical intoxication symptoms include

- euphoria,
- nausea,
- drowsiness (nodding off),
- constipation,
- pupil constriction,
- impairments in attention, memory, and social functioning.

High doses of opioids may cause an initial euphoria followed by apathy, dysphoria, slurred speech, respiratory depression, decreased heart rate, and possible coma or death. Risk of death by overdose and respiratory depression is a major concern, especially when opioids are combined with antihistamines, sedatives (e.g. barbiturates, benzodiazepines) or alcohol. For example, promethazine (e.g. Phenergan) is a popular antimimetic prescription drug that can be mixed with codeine or oxycodone for recreational use and referred to as sizzurp.

Acuity of intoxication is partially dependent on the route of administration. For example, rapid release oxycodone (OxyContin) in a crushed form provides a more intense intoxication than extended release capsules that are metabolized in the digestive tract. Injecting or snorting crushed pills produces a quick flood of neurotransmitters in the DA reward pathway system, whereas opioids that are gradually released (i.e. transdermally administered) closely resemble endogenous processes in the brain.

Symptoms of opioid withdrawal generally follow the opposite course of symptoms of opioid intoxication. Acute opioid withdrawal symptoms can occur as soon as 2.5 h after intake and last up to 4 days. Withdrawal symptoms resemble influenza and include

- insomnia,
- dysphoric mood,
- nausea and vomiting,
- diarrhea,
- muscle pain and aches,
- yawning,
- lacrimation or rhinorrhea (i.e. runny eyes or nose),
- pupil dilation, piloerection (i.e. goose bumps), or sweating,
- fever,
- involuntary leg movements.

Residual withdrawal symptoms following heavy opioid use, such as anhedonia, anxiety, insomnia, and drug craving can continue for months after cessation. Withdrawal symptoms are a main cause of relapse, as many individuals return to opioid use in an attempt to relieve or prevent uncomfortable withdrawal symptoms.

Treatment

The combination of psychotherapy and pharmacotherapy is most effective for the treatment of opioid use disorders. Behavioral therapies can help strengthen coping skills, enhance motivation for change, improve social functioning, address maladaptive cognitions and behaviors, and assist patients in building a substance-free life, and social support system. A range of behavioral models may be used, including cognitive-behavioral, harm-reduction, voucher-based contingency management, relapse prevention, and traditional, abstinence-based 12-step models. Pharmacotherapies work in conjunction with behavioral therapies to ameliorate the physiological symptoms of withdrawal and craving, thereby helping to prevent relapse.

The first step of treatment is often opioid detoxification, which may occur in a medically monitored facility or under physician supervision in an office-based setting, depending on the level of addiction. Several treatment options exist for opioid detoxification. Opioid replacement therapies substitute an opioid agonist (e.g. methadone, buprenorphine) and slowly taper the dose until the patient no longer requires any opioid agents. A second option is to use a nonopioid agent (e.g. clonidine) to decrease elevated noradrenaline (NE) levels present during withdrawal, which mitigates withdrawal symptoms. A third option is to shorten the length of the opioid withdrawal syndrome via a rapid or ultrarapid detoxification process. Withdrawal is induced and accelerated by administering an opioid antagonist to quickly displace the remaining drug from the receptor. Nonopioidergic drugs are typically given to treat subsequent withdrawal symptoms. This relatively new form of opioid detoxification is controversial in terms of safety and long-term efficacy. Critics argue that the procedure is unsafe due to the potential for cardiorespiratory complications and aspiration while under anesthesia.

Medication-assisted treatments for opioid dependence consist of opioid replacement therapies, opioid antagonists, and mixed opioid agonist–antagonists. LAAM and methadone maintenance are the two full agonist replacement therapies approved by the FDA in the United States. Low doses of these opioid agonists are administered to patients at regularly scheduled intervals in conjunction with psychotherapy. Of these two replacement therapies, methadone has been the preferred treatment since 2001 when the FDA

administered a warning that LAAM may cause cardiac arrhythmia. LAAM has since been withdrawn from the market. For further information regarding methadone maintenance as a treatment for opioid dependence, *see* Methadone Maintenance.

Buprenorphine (Subutex®) is a partial agonist commonly used in the treatment of opioid dependence. It works as a partial agonist at the mu receptor while simultaneously serving as an antagonist at the kappa receptor. Since buprenorphine does not fully activate the mu receptor, it can precipitate withdrawal among individuals who chronically abuse high doses of opioids. Buprenorphine combined with naloxone, an opiate antagonist (Suboxone®) is frequently used to prevent misuse (e.g. injection of the drug to get high or selling it on the street) and to treat opioid dependence. Antagonists such as naltrexone or nalmefene can also be useful in the treatment of opioid dependence, as they have low abuse liability. For more information regarding the use of antagonists for the treatment of opioid dependence, *see* Antagonists for the Treatment of Opioid Dependence.

Central Nervous System Depressants

Medical Uses

The central nervous system (CNS) depressant drug class is frequently referred to as the sedative-hypnotic drug class and includes a wide spectrum of prescription medications used to induce anesthesia and treat anxiety, sleep, and seizure disorders. Barbiturates, benzodiazepines, hypnotics, and tranquilizers are included in this group. Barbiturates are approved for use as anticonvulsants, hypnotics, and sedatives. Benzodiazepines are used to treat anxiety disorders and alcohol withdrawal. When taken as prescribed at low to moderate doses, benzodiazepines offer many of the same beneficial effects as other sedative-hypnotics without the severe side effects and overdose potential. However, dependence on benzodiazepines can develop after chronic use. Dependence occurs primarily in populations with preexisting vulnerabilities to substance use disorders and is more likely evidenced by physiological withdrawal symptoms following cessation of use (i.e. physiological dependence) rather than positive reinforcing properties obtained during use. Compared to other medications with abuse potential (e.g. opioids and amphetamines), benzodiazepines have a relatively safe and low abuse profile.

Neurobiology

CNS depressants increase the inhibitory neurotransmitter GABA, which serves to decrease neuronal firing. GABA and benzodiazepine receptors are located in most synapses throughout the brain. Due to the pervasive expression of these receptors, it is difficult to isolate specific neural circuits activated by benzodiazepine and GABAergic agents. However, specific symptoms of intoxication and withdrawal are modulated by multiple neurotransmitter systems acting within different brain structures. For example, increased inhibition in the Papez circuit of the limbic system and basolateral amygdala may be responsible for the anxiolytic response during drug use and subsequent anxiogenic response during withdrawal. Barbiturates are believed to have a particularly strong effect on the mesencephalic reticular activating system. If synapses are activated postsynaptically by barbiturates, their receptors may be found on neurons in the cuneate nucleus, substantia nigra, and thalamus. Presynaptic activation of these agents occurs in the spinal cord. The most prominent neurotransmitter systems activated during CNS depressant use are GABA and glutamate. Some CNS depressants, such as glutethimide (Doriden), produce anticholinergic effects as well.

Pharmacology

GABA receptors are located in the majority of the synapses in the CNS and play an important role in neuronal firing (e.g. seizure disorders), regulating emotion, and controlling cognition, pain, sleep, and motor function. GABA has three receptor types: GABAa, GABAb, and GABAc. Benzodiazepines activate GABAa receptors, which are primarily responsible for the anxiolytic effects at low doses, sedation at moderate doses, anesthetic properties at high doses, and fatality in the case of overdose. Different GABAa receptor subtypes are responsible for different drug effects, meaning that one receptor subtype is responsible for the sedating effects of a drug while another produces anxiolytic effects.

A glutamatergic hypothesis has been proposed, suggesting that augmented GABA inhibitory neurotransmission produces a compensatory increased glutamatergic excitatory response. Sudden withdrawal of benzodiazepines reduces GABAergic neurotransmission while there is still a residual enhanced glutamatergic response remaining. This excess glutamate may be responsible for the anxiogenic and epileptic symptoms possible during benzodiazepine withdrawal.

Benzodiazepines also have their own class of receptors in the brain, which is subdivided into three receptors types: BZ1, BZ2, and BZ3. A close relationship exists between GABA and BZ receptors, such that activation of BZ receptors can in turn produce a stronger inhibitory response from GABAa receptors. Recent studies have shown that drugs classified as nonbenzodiazepine hypnotics activate the same receptors as benzodiazepines, despite significant differences in their chemical structure.

Barbiturates activate GABAergic synapses in addition to increasing the binding affinity of benzodiazepine and GABA receptors. GABAa receptors in particular appear to be the primary site of augmented inhibitory action for barbiturates. In addition, barbiturates block the glutamate subreceptor α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA). Inhibiting the AMPA subreceptor of this excitatory neurotransmitter at least partially accounts for the anticonvulsant property of barbiturates.

The primary difference between barbiturates and benzodiazepines is the mechanism by which they open ion channels. Barbiturates cause the channel to open for a longer period of time, whereas benzodiazepines cause the channel to open more frequently. This difference in phamacodynamics partially explains why barbiturates are more likely to cause an overdose. For further information regarding the neuropharmacology of benzodiazepines, *see* Neuropharmacology of Benzodiazepines.

Intoxication and Withdrawal Syndrome

Symptoms of CNS depressant intoxication include

- decreased anxiety,
- subjective feelings of well-being and confidence,
- lowered inhibitions,
- decreased heart rate and blood pressure,
- depressed respiration (arrested respiration at high doses),
- poor concentration,
- confusion,
- dizziness,
- sedation and fatigue,
- impaired coordination, memory, and judgment.

While under the influence of CNS depressants, particularly benzodiazepines, learning is impaired and memory loss for events that took place while intoxicated may occur. Combining CNS depressants with alcohol or another sedating substance increases the risk of loss of consciousness and fatalities. Two drugs from this class, γ-hydroxybutyrate (GHB) and flunitrazepam (Rohypnol), are especially dangerous. Since GHB and flunitrazepam (also known as date-rape drugs) are colorless, odorless, and can be easily dissolved into liquids, they are used by sexual perpetrators. Once ingested, these drugs act quickly to induce memory loss. Withdrawal symptoms from CNS depressants include

- anxiety,
- agitation,
- tremors,
- nightmares,
- insomnia,
- anorexia,
- nausea and vomiting,
- orthostatic hypotension, and
- fever.

If left untreated, withdrawal symptoms may escalate to delirium, disorientation, hallucinations, and seizures. Withdrawal from CNS depressants may be life threatening if doses are not appropriately tapered. Medically supervised detoxification is necessary. The onset of withdrawal symptoms for short-acting CNS depressants (e.g. pentobarbital, meprobomate, methaqualone) occurs 12–24 h after the last dose, while withdrawal from longer-acting CNS depressants (e.g. phenobarbital, diazepam) does not begin until 24–48 h after the last dose.

Treatment

Detoxification from CNS depressants should be medically monitored due to the elevated risk of seizures. Chronic use of high doses of benzodiazepines or other depressants increases the risk for a severe withdrawal syndrome. In order to minimize the severity of withdrawal symptoms, doses of benzodiazepines or pentobarbital can be gradually tapered. Long-acting barbiturates, most commonly phenobarbital, can be substituted for short-acting agents and the long-acting medications are slowly withdrawn. In addition, anticonvulsants such as carbamazepine or valproate can be substituted to treat withdrawal symptoms.

Aside from available detoxification strategies and pharmacologic treatment for withdrawal symptoms, there are no identified evidence-based psychotherapies specifically targeting the treatment of CNS depressant use disorders. For a discussion of evidence-based practices for treating generalized addictive disorders and further information regarding pharmacotherapies for sedative dependence, *see* Evidence-Based Treatment and Medications for sedative dependence.

Stimulants

Medical Uses

Included in the stimulant class are amphetamines, analeptics, anorexiants (also referred to as anorectics), and methylphenidates. Of these, the most commonly abused stimulants are amphetamines and methylphenidates. Amphetamines were originally developed as a diet pill to decrease appetite and increase the breakdown of adipose tissue. Amphetamines have since been approved for the treatment of narcolepsy and ADHD. Methylphenidate and dexmethylphenidate have also been approved for the treatment of ADHD. Methylphenidate has an additional approved clinical indication for the treatment of narcolepsy.

The analeptics are most often misused for their cognitive enhancing properties. Modafinil and armodafinil are approved for the treatment of narcolepsy, obstructive

sleep apnea-hypopnea syndrome, and shift-work sleep disorder. Doxapram is another analeptic used to stimulate respiration following drug overdose, postoperative anesthesia, and chronic obstructive pulmonary disease. The anorexiants are approved as a short-term treatment for obesity when combined with caloric restriction. Many of the drugs in this class (e.g. sibutramine, fenfluramine, phenmetrazine, phentermine) have been discontinued in the United States due to adverse cardiovascular side effects and concerns about abuse potential. For information about methamphetamine, *see* Methamphetamine Addiction.

Neurobiology

The therapeutic effects of stimulants stem from increased activation in the prefrontal cortex, limbic structures, and hypothalamus. Among other functions, the hypothalamus is important for regulating stress, arousal, feeding, blood pressure, and autonomic activity. Imaging studies have revealed that individuals with narcolepsy have structural abnormalities in the hypothalamus, whose interaction with other limbic structures such as the cingulate cortex may be responsible for cataplexy, a predominant symptom of narcolepsy. In addition, the hypothalamus and its functional connectivity to limbic and cortical structures are important in controlling food intake. The hedonic component of obesity has been linked to limbic and cortical structures, such as the nucleus accumbens, ventral tegmental area, orbitofrontal cortex, and amygdala.

ADHD correlates with deficits in frontocortical areas, specifically in structures responsible for impulsivity, attention, and motor activity. Likewise, brain imaging studies have revealed that methylphenidate, a popular ADHD pharmacotherapy, increases activity in the orbital and medial prefrontal cortices. Increased levels of dopaminergic and noradrenergic neurotransmission within cortical, limbic, and hypothalamic structures are likely responsible for mitigating symptoms in ADHD, obesity, and narcolepsy. The intensity of the hedonic effects of stimulants is related to the increase of extracellular DA in mesocorticolimbic pathways (i.e. reward pathways of the brain).

Pharmacology

Stimulants work by activating the sympathetic nervous system, primarily through increasing the levels of DA and NE available in the synapse. When taken as prescribed, stimulants slowly increase neurotransmitters in a manner that mimics endogenous processes. However, when taken in excess or through alternative routes of administration (e.g. crushing and snorting), DA and NE flood the synapses and produce symptoms of intoxication. This rapid increase is caused by higher levels of DA and NE being released into the synapse while reuptake of these neurotransmitters is decreased. Within frontocortical structures, DA functions as an inhibitory neurotransmitter, so stimulant pharmacotherapies enhancing DA release decrease impulsivity, increase attention, and improve vigilance. Anorexiants such as sibutramine increase serotonin levels in addition to DA and NE.

The pharmacological means by which armodafinil and modafinil operate are unclear. Although evidence regarding modafinil's influence on dopaminergic and noradrenergic neurotransmission remains under study, it appears that some level of interaction exists. One hypothesis is that modafinil acts through DA release to stimulate adrenergic receptors. However, the exact mechanism of action is difficult to identify, as modafinil appears to interact with dopaminergic, serotonergic, noradrenergic, GABAergic, orexinergic, histaminergic, and glutamatergic neurotransmission on varying levels.

Intoxication and Withdrawal Syndrome

Prescription stimulant medications have the same mechanism of action as other psychostimulants (e.g. cocaine) and their symptoms of intoxication closely resemble those of other stimulants. Intoxication symptoms include

- enhanced confidence and feelings of well-being,
- increased heart rate, blood pressure, respiration, and body temperature,
- constricted blood vessels,
- dilated bronchioles,
- increased arousal, alertness, and energy,
- euphoria,
- decreased appetite, and
- elevated blood sugar.

Much like cocaine, amphetamines can induce psychosis, including symptoms of paranoia, hallucinations, and delusions. Chronic, long-term use of stimulants can produce stereotyped, repetitive behaviors due to continual locomotor stimulation. Extremely high doses can induce cardiac arrest. Another subtle but dangerous effect of amphetamines is the increase in body temperature. When used to enhance physical performance, the increased body temperature combined with extreme exertion can be fatal. Tolerance to certain effects of stimulants, such as suppression of appetite, develops rapidly after consistent predictable intervals of use. Dangers of overdose are especially high when these drugs are taken via alternative routes (e.g. smoking, injecting, crushing, and snorting). In addition, the negative side effects of stimulants can be masked when they are combined with opioids such as heroin, rendering speed balling a hazardous combination.

Adhering to a routine daily schedule of prescribed stimulant medication dramatically decreases the risk of

addiction. Relatively few individuals who are prescribed stimulant medications for ADHD develop abuse or dependence. Although dependence on ADHD medications is a significant concern, some research shows that the likelihood of developing a future substance use disorder is lower for individuals treated pharmacologically for ADHD compared with individuals with untreated ADHD.

Although withdrawal from stimulants is not life threatening, symptoms of the crash following the high are often so dysphoric that it results in relapse. Anhedonia and craving are particularly strong after chronic stimulant use cessation. In addition to anhedonia, symptoms of withdrawal include

- excessive sleep or insomnia,
- vivid, disturbing dreams,
- depressive symptoms,
- psychomotor retardation or agitation, and
- increased appetite.

Treatment

Pharmacological treatments for amphetamines and other stimulants target symptoms of withdrawal, stimulant-induced disorders, or co-occurring psychiatric disorders. Similarly, detoxification from stimulants consists primarily of supportive measures, ensuring that the person is in a safe environment and treating individual withdrawal symptoms. Medication development for amphetamine dependence is described in Medication Development for Amphetamine Dependence.

With regard to behavioral therapies, contingency contracting is frequently used in the treatment of stimulant dependence. This operant conditioning intervention is based on a mutually agreed contract between clinician and client, stating that the client agrees to perform certain behaviors. Failure to perform these behaviors may result in aversive consequences, whereas successful performance leads to rewards. Urine drug screens are often an important component of this intervention. Prize incentives contingency management for substance abuse treatment is an evidence-based practice utilizing contingency contracting to treat individuals with stimulant and cocaine dependence. Participants draw chips from a bowl designating they have won a prize valued between $1 and $100. As patients perform more desirable behaviors (e.g. negative urine drug screen tests), the number of chips they are allowed to draw from the bowl increases. The matrix model is another evidence-based practice approved for the treatment of stimulant abuse and dependence. The intervention uses group and individual therapy sessions to provide psychoeducation, teach relapse prevention skills, enhance self-esteem, and offer social support.

Over-the-Counter Medications

The nonmedical use of OTC medications is becoming increasingly popular among adolescents and young adults, in particular, due to the accessibility, perceived safety, legality, lowered social stigma, and the financial affordability of these drugs. Slang terms such as pharming (i.e. the selling and use of OTC medications for recreational purposes) and sheeting (i.e. measuring use in terms of number of full cards of medication consumed rather than number of gel capsules per card) have arisen to describe practices within this subculture. Minimal research is available on the abuse and dependence syndromes resulting from chronic use of OTC medications.

Medications containing ephedrine and pseudoephedrine were initially released as OTCs but are now behind-the-counter (BTC) medications. As production of methamphetamine from ephedrine and pseudoephedrine surged in the mid-2000s, the Drug Enforcement Administration (DEA) changed federal legislation regarding the distribution of these products in an effort to reduce drug diversion. Regulation of ephedrine and pseudoephedrine sales varies by state; however, federal law requires that anyone purchasing ephedrine or pseudoephedrine must present photo identification in addition to providing a name, address, date/time of purchase, and signature to the pharmacist. Some states refuse to sell ephedrine-containing products except within inpatient hospital settings.

Ephedrine is currently FDA-approved for the treatment of asthma. This medication, also known as herbal X-tacy, was initially sold as a weight-loss enhancer and illicitly distributed as a safe ecstasy substitute. Unlike amphetamines, ephedrine does not cross the blood–brain barrier as easily, so it produces fewer subjective symptoms except when abused at high doses. However, the stimulant effects of ephedrine on the cardiovascular system are problematic and misuse has led to fatalities, especially when combined with monoamine oxidase inhibitors. Pseudoephedrine is a nasal and sinus decongestant that increases release of noradrenaline and constricts blood vessels in the nasal passage. Commonly known as Sudafed®, it also has limited subjective effects except when taken at high doses. Because of the regulation of ephedrine and pseudoephedrine, pharmaceutical companies have begun substituting phenylephrine for pseudoephedrine. Although phenylephrine may produce giddiness and restlessness at high doses, it cannot be diverted for methamphetamine production. Currently, many phenylephrine and pseudoephedrine products share the same brand names (refer to Table 76.1); however, products containing phenylephrine as the active ingredient are available OTC, whereas pseudoephedrine-containing products are available BTC.

DXM is an antitussive/expectorant combination contained in over 140 OTCs. When taken as prescribed, it is a relatively safe drug with few side effects. It suppresses the cough center in the medulla and serves as an agonist to sigma receptors. This poorly understood receptor class was originally categorized as an opioid receptor; however, its structural and functional differences have now designated it as its own class. DXM's actions on sigma receptors may contribute to the drug's dissociative and hallucinogenic properties at high doses. At low doses, DXM serves as a glutamate antagonist. Taken in significantly higher than recommended doses, it serves as an *N*-methyl-D-aspartate (NDMA) antagonist and serotonin agonist. Although this drug has minimal side effects when taken as prescribed, there have been multiple reports of fatalities among young adults due to overdose of DXM. Symptoms of intoxication for DXM include stupor, hyperexcitability, and euphoria. At high doses, DXM may produce nausea, vomiting, pupil dilation, hallucinations, delusions, seizures, paranoia, irregular heartbeat, hyperthermia, coma, and death. Combining DXM with sedatives, selective serotonin reuptake inhibitors (SSRIs), or antihistamines is especially dangerous.

Antihistamines serve as antagonists to histamines at the H1 receptor, but also interact with dopaminergic, serotonergic, noradrenergic, opioidergic, and acetylcholinergic systems. They are usually available OTC in low doses. Antihistamines are often abused in conjunction with other drugs to produce an augmented high. For example, tripelennamine is often combined with opiates, particularly pentazocine (in a mixture referred to as T's & Blues) to produce a euphoric rush similar to heroin. Diphenhydramine is contained in a number of cold and flu medications including Benadryl, Theraflu strips, and Midol, which are frequently used for recreational purposes in conjunction with alcohol or other sedatives. Dimenhydrinate, commonly marketed as Dramamine®, is sometimes abused for its euphoric and hallucinogenic properties. Other antihistamines include doxylamine (e.g. Nyquil, Alka-Seltzer), Nytol, and Compoz. Many of these drugs contain alcohol and intensify the effects of alcohol or other sedatives when used in combination with other drugs.

SUMMARY

Over the past decade, the nonmedical use of prescription and OTC medications has increased significantly and is a serious public health concern. In contrast to other commonly abused substances in our society, prescription and OTC medications are unique in that their consumption is endorsed by health care professionals, they serve as a primary treatment for certain medical conditions, and they represent a legitimate means of improving physical health for many patients. Because of this some individuals, in particular teenagers and young adults, develop a false sense of safety regarding the nonmedical use of prescription and OTC medications. Early treatment typically involves detoxification and the management of withdrawal symptoms, a common cause of relapse. While further research on the development of evidenced-based treatments for nonmedical prescription and OTC medication use is clearly needed, a combination of pharmacotherapy and psychotherapy is recommended.

Acknowledgment

This work was supported in part by grant K23 DA021228 (SEB) from the National Institute on Drug Abuse.

SEE ALSO

Methamphetamine Addiction

Glossary

Abuse a maladaptive pattern of use leading to clinically significant impairment or distress as evidenced by one or more DSM-IV diagnostic criteria in the past 12 months (e.g. recurrent failure to fulfill major obligations at work, school, or home; recurrent use in situations in which it is physically hazardous; continued use despite having recurrent social or interpersonal problems caused or exacerbated by the use).

ADHD attention deficit hyperactivity disorder.

AMPA α-Amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid.

Antagonists drugs that bind to receptors but do not activate them, thereby blocking the receptor and preventing activation from other chemicals.

Behind-the-counter medications (BTCs) US FDA-approved medications that require photo identification, contact information, and signature in order to purchase.

DA dopamine.

Dependence often used synonymously with addiction. A maladaptive pattern of use that indicates a loss of control and clinically significant impairment or distress, as evidenced by three or more DSM-IV diagnostic criteria in the past 12 months (e.g. tolerance, withdrawal symptoms, persistent desire, or unsuccessful efforts to cut down or quit, taking more of the substance than was planned).

DXM dextromethorphan

Evidence-based practice (EBP) an intervention tested and accepted by consensus or expert review and demonstrated to yield a specific clinical outcome (i.e. objective, concrete, measurable change in client status follows implementation of practice).

FDA US Food and Drug Administration

Full agonists drugs that bind to receptors and fully activate them, thereby closely mimicking endogenous processes.

GABA γ-aminobutyric acid.

GHB γ-hydroxybutyrate

LAAM levacetylmethadol.

LSD lysergic acid diethylamide.

Misuse incorrect use of a medication (e.g. taking more than was prescribed, taking it more often than was prescribed, using it for a purpose other than that for which it was prescribed).

NAc nucleus accumbens.

NE noradrenaline.

Nonmedical use the use of a medication that was not prescribed by a physician or was used only for the experience or feeling it caused.

Over-the-counter medications (OTCs) US FDA-approved medications available directly to consumers without a prescription.

Partial agonists drugs that bind to receptors but do not activate them with the same intensity as endogenous chemicals or full agonists.

PCP phencyclidine.

PFC prefrontral cortex.

Pharming a slang term used to describe the distribution and use of prescription and OTC medications for recreational purposes.

Physiological dependence tolerance for a medication, emergence of withdrawal symptoms upon cessation of the medication, or continued use of the medication to prevent withdrawal symptoms. Note that a person could have physiological dependence on a medication but not be addicted to it (see definition for Dependence).

Prescription medications US FDA-approved medications that require authorization (i.e. a prescription) from a licensed medical professional in order to purchase.

Pseudoaddiction Drug-seeking and other behaviors that are phenotypically consistent with psychological dependence but actually result from insufficient treatment (e.g. lack of pain relief). Once the condition is adequately treated, the drug-seeking and other related behaviors cease.

VTA ventral tegmental area.

Further Reading

Allison, C., Pratt, J., 2003. Neuroadaptive processes in GABAergic and glutamatergic systems in benzodiazepine dependence. Pharmacology and Therapeutics 98, 171–195.

AmeraChem Inc, 2010. Drug Identification Bible. AmeraChem Inc, Grand Junction, CO.

Brady, K., Back, S., Greenfield, S. (Eds.), 2009. Women and Addiction: A Comprehensive Handbook. Guilford Press, New York.

Ford, J., 2009. Misuse of over-the-counter cough or cold medications among adolescents: prevalence and correlates in a national sample. Journal of Adolescent Health 49, 505–507.

Halpert, A., Olmstead, M., Beninger, R., 2002. Mechanisms and abuse liability of the anti-histamine dimenhydrinate. Neuroscience and Biobehavioral Reviews 26, 61–67.

Koob, G., Volkow, N., 2010. Neurocircuitry of addiction. Neuropsychopharmacology Reviews 35, 217–238.

Koob, G., Wee, S., 2010. The role of the dynorphin–κ opioid system in the reinforcing effects of drugs of abuse. Psychopharmacology 210, 121–135.

Kuhn, C., Swartzwelder, S., Wilson, W., 2008. Buzzed: The Straight Facts About the Most Used and Abused Drugs from Alcohol to Ecstasy. W.W. Norton, New York.

Levine, D., 2007. 'Pharming': the abuse of prescription and over-the-counter drugs in teens. Current Opinion in Pediatrics 19, 270–274.

Pliszka, S., 2007. Pharmacologic treatment of attention-deficit/hyperactivity disorder: efficacy, safety and mechanisms of action. Neuropsychology Review 17, 61–72.

Schatzberg, A., Nemeroff, C. (Eds.), 2006. Essentials of Clinical Psychopharmacology, second ed. American Psychiatric Publishing, Arlington, VA.

Shippenberg, T., Vries, T., 2002. Neural systems underlying opiate addiction. Journal of Neuroscience 22, 3321–3325.

Ruiz, P., Strain, E., Langrod, J. (Eds.), 2007. The Substance Abuse Handbook. Lippincott Williams & Wilkins, Philadelphia, PA.

American Psychiatric Association, 2000. Diagnostic and statistical manual of mental disorders, fourth ed. (text revision). American Psychiatric Association, Washington, DC.

Relevant Websites

http://www.thecochranelibrary.com – The Cochrane Library.

http://www.accessdata.fda.gov/scripts/cder/drugsatfda/ – Drugs@FDA: FDA drug approved products.

http://www.factsandcomparisons.com/ – Facts & Comparisons (Wolters Kluwer Health).

https://nsduhweb.rti.org/ – National Survey on Drug Use and Health (NSDUH).

http://www.nida.nih.gov – National Institute of Drug Abuse (NIDA).

http://www.drugabuse.gov/ResearchReports/Prescription/Prescription.html – NIDA Research Report Series: Prescription Drugs Abuse and Addiction (2005).

http://www.samhsa.gov/ – Substance Abuse and Mental Health Services Administration (SAMHSA).

http://nrepp.samhsa.gov/ – SAMHSA's National Registry of Evidence-based Programs and Practices (NREPP).

http://www.justice.gov/dea/index.htm – US Drug Enforcement Administration.

CHAPTER

77

Tobacco

Angela Attwood[*], *Paul Aveyard*[$], *Linda Bauld*[**], *John Britton*[§], *Peter Hajek*[‖], *Gerard Hastings*[#], *Ann McNeill*[§], *Marcus Munafò*[*], *Rachael Murray*[§], *Lion Shahab*[¶], *Robert West*[¶]

[*]School of Experimental Psychology, University of Bristol, Bristol, UK
[$]Primary Care Clinical Sciences, University of Birmingham, Edgbaston, UK
[**]School of Management, University of Stirling, Stirling, UK [§]Division of Epidemiology and Public Health, Clinical Sciences Building, Nottingham City Hospital, Nottingham, UK [‖]Tobacco Dependence Research Unit, Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, Whitechapel, London, UK [#]Institute for Social Marketing, University of Stirling, Stirling, UK
[¶]Health Behaviour Research Centre, University College London, London, UK

OUTLINE

TOBACCO: AN INTRODUCTION

The use of tobacco dates back thousands of years. It is cultivated from two main species of tobacco plant (*Nicotiana tabacum* and *Nicotiana rustica*) originating in the Americas and West Indies. There are numerous ways in which tobacco can be consumed, and there are important differences between smoked and smokeless tobacco, and between different forms of delivery device. However, the focus here is on smoked tobacco (e.g. cigarette) as this is the most common form of tobacco consumption and is associated with the greatest degree of dependence and ill-health. Nicotine is the primary psychoactive constituent of tobacco, and is thought to underlie many of the behavioral and addictive properties of tobacco use. However, nicotine is only one of around 4000 compounds that are released when tobacco is burned. Therefore, although nicotine is considered to be the most important constituent of tobacco, it does not underlie all of the effects of smoking and cannot fully explain the smoking phenomenon. For example, smoking is more widely abused than other methods of nicotine delivery (e.g. chewed tobacco), and nicotine replacement therapies (NRTs) have limited efficacy in

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00077-2

helping people to quit smoking suggesting that nicotine is not the only factor important in smoking dependence. In addition, the main adverse long-term effects of smoking, including respiratory and cardiovascular disease are predominantly related to non-nicotinic compounds in tobacco, such as carbon monoxide and tars.

When tobacco is smoked, nicotine is taken into the lungs and is quickly absorbed into the bloodstream and distributed throughout the body. It readily crosses the blood brain barrier, first entering the brain around 10–15 s after inhalation. Given the speed of activation, smoked tobacco is the most popular vehicle of recreational nicotine administration. However, nicotine is effectively absorbed via other areas of the body, including through membranes in the mouth (cheeks and gums), nasal membrane, and through the skin (transdermal), all of which have been used as sites of absorption for NRTs.

Nicotine is primarily metabolized by the liver (around 70–90%) and excreted in urine. A special enzyme called Cytochrome P450 2A6 (CYP2A6) transforms nicotine into its principal metabolite, cotinine. The elimination half-life is approximately 2 h, although this can vary depending on several factors such as the frequency and heaviness of smoking and individual differences in metabolism (e.g. degree of CYP2A6 activity).

Nicotine binds to nicotinic acetylcholine receptors both in the central nervous system and the periphery. It acts largely as a stimulant drug, increasing heart rate and blood pressure and initiating release of epinephrine (adrenaline) from the adrenal glands. Stimulation of these receptors in the brain has been associated with increased psychomotor activation and improved cognitive function, particularly working memory and sustained attention, although these effects appear to vary across individuals and are also task dependent, with generally better outcomes being observed on relatively simple mental performance tasks rather than complex problem solving. Nicotine administration has also been shown to increase calmness and reduce tension in abstinent smokers, often leading to the belief that smoking is anxiolytic. Somewhat paradoxically to this, smokers generally display higher levels of anxiety than non-smokers, indicating that the anxiolytic effects associated with abstinence may be a reversal of withdrawal-related tension. This is supported by the fact that nicotine is a stimulant (rather than sedative) drug that often induces feelings of heightened arousal and tension when administered acutely to non-smokers.

Importantly, like the majority of recreationally abused drugs, nicotine also stimulates activation of the reward pathways in the brain, which accounts for at least some of the rewarding properties of smoking. In higher doses, nicotine is toxic and can induce a plethora of aversive symptoms including vomiting, dizziness, excessive sweating and salivation, tremors, and seizures. However, cases of nicotine overdose are extremely rare, and tend to occur through non-volitional intake rather than excessive smoking (e.g. accidental swallowing or exposure to high levels of nicotine when harvesting tobacco). Early use of nicotine can induce nausea and vomiting, possibly due to activation within the area postrema (the area of the brain that controls vomiting), although tolerance to this effect occurs quickly with repeated use.

THEORY OF ADDICTION

The process by which smokers become addicted to tobacco is quite well understood. Nicotine plays a pivotal role. Cigarettes are the dominant and most addictive form of tobacco. When the smoke is inhaled nicotine is absorbed rapidly via the large surface area of the lungs and transported to the brain where it attaches to 'nicotinic acetylcholine receptors'. This has different effects in different parts of the brain but one area that is of particular importance is a part of the midbrain called the 'ventral tegmental area (VTA)'. When nicotine attaches to receptors in the VTA it leads to an increased rate of firing of neurones originating in this area. Some of these neurones project to a part of the forebrain called the 'nucleus accumbens (NAcc)'. The nerve terminals in this area release another neurotransmitter called dopamine. This dopamine release causes the smoker's brain to pay attention to what activity it had just performed and the situation, and strengthen the impulse to engage in that activity when the same situation arises. This does not require conscious awareness and in that sense is 'automatic'. Thus the smoker experiences urges to smoke in situations where smoking has occurred in the past and these can be very strong.

Repeated exposure to nicotine in this way also has a chronic effect on the brain which leads smokers to experience a kind of 'nicotine hunger' – a drive to smoke – when nicotine concentrations in the brain are depleted. Thus even in situations where smoking has not occurred, smokers may feel a need to smoke.

In addition to this, adaptation to nicotine leads the brain to experience a range of adverse withdrawal symptoms when nicotine concentrations become depleted. These include: irritability, depressed mood, restlessness, difficulty concentrating, sleep disturbance, and increased appetite. There is also usually rapid weight gain. This has two effects. One is that the adverse experiences create a need to smoke to relieve them. The second is that the repeated experience of mood disturbance and relief by smoking creates a strong expectation that smoking

helps with stress, regardless of how that stress is generated. Thus when smokers experience a stressful situation which may be months or years after they have stopped smoking, they will often feel a need to smoke.

In summary, nicotine from tobacco (and particularly when delivered rapidly via cigarettes) establishes (1) urges to smoke in situations where smoking has occurred in the past, (2) a drive to smoke when brain nicotine concentrations are depleted, and (3) a need to smoke to avoid unpleasant mood and physical symptoms together with an expectation that smoking will help with stress.

There is considerable variation between individuals and over time in susceptibility to these effects. Moreover, there is reason to believe that a range of other factors can exacerbate or mitigate the development and expression of dependence. Thus smokers can develop a strong 'smoker identity' to which they are attached and the sense of belonging to a social group of smokers can motivate continued smoking. It has also recently been hypothesized that the effects of nicotine may be dependent on, or amplified by, the presence of monoamine oxidase inhibitors (MAOIs) which are present in cigarette smoke. Most smokers report that they enjoy smoking but this does not appear to play a role in leading them back to smoking if they try to stop.

EPIDEMIOLOGY OF SMOKING

Tobacco use prevalence has steadily decreased since the middle of the last century in the developed world, standing at around 20–30% in many Western countries, but is stagnating or still increasing in the developing world; the vast majority of smokers now live in developing nations or nations in transition. Consequently, the worldwide population of smokers is estimated to reach 1.7–1.9 billion by 2020, more than ever before.

Tobacco use is initiated at a young age; most smokers will have started before their 18th birthday. Adolescent smoking rates vary considerably from country to country but unlike adult smoking rates, there is currently no clear trend for changes in adolescent smoking prevalence. Smoking rates also vary with age; while smoking is less common in both young and old age, prevalence peaks somewhere before middle age. This is primarily due to two factors; less dependent smokers start to give up smoking as they grow older, while more dependent smokers continue to smoke and as a consequence start to die prematurely from middle age onwards, resulting in a bell-shaped distribution of smoking prevalence across the human life span.

Globally, smoking is more common among men than women but this gender gap has much reduced in the developed world, having virtually disappeared in countries like Australia. Whilst female smoking prevalence is still much lower in developing countries, this is likely to change with shifts in social and cultural conventions and targeted marketing by tobacco companies.

Besides gender, there exist a number of other important socio-demographic determinants of smoking such as ethnicity. Taking the United Kingdom as an example, smoking prevalence among minority ethnic groups as a whole is somewhat lower than among the general population but much higher in some ethnic minorities, notably those originating from the Indian subcontinent. Moreover, gender biases for smoking also persist in ethnic minority groups; for example, in the Asian community in the United States smoking prevalence among men is three times higher, and in the Hispanic community two times higher, than among women.

One of the strongest determinants of tobacco use is socio-economic status (SES), but the direction of this relationship is variable. In some countries, usually those with weak tobacco control measures and less public awareness of the consequences of smoking, cigarette consumption is still mostly determined by income, leading to a positive correlation between smoking and SES. However, in countries where awareness of the negative impact of smoking is high, there is a negative relationship between smoking and SES. It appears that better-off smokers are not more likely to attempt to stop, but are more likely to succeed in giving up smoking, possibly due to deprived smokers lacking adequate financial, emotional, and environmental resources to aid smoking cessation.

PUBLIC HEALTH ISSUES

Tobacco smoking is the largest avoidable cause of death, disability, and social inequality in life expectancy in the developed world, and with the growing epidemic of smoking throughout the developing world is set to become the biggest avoidable cause of disease globally. Smoking killed about 5 million people in 2006, and is projected to kill about 10 million in 2025. Although smoking prevalence has been declining in many developed countries over recent decades, the reverse is true in many other countries. As a result, the majority of the burden of tobacco-related deaths over the coming decades will occur in low- and middle-income countries. Where smoking is currently prevalent, smoking typically accounts for up to one in five of all deaths.

Half of all regular cigarette smokers die from a disease caused by their smoking, typically losing about 10 years of life. Smoking causes over 50 fatal illnesses and a range of chronic diseases, but the majority of deaths from smoking arise from three causes: cardiovascular disease,

lung cancer, and chronic obstructive pulmonary disease (COPD); the majority of deaths from lung cancer, and of chronic obstructive disease in the developed world, are due to smoking. Globally, smoking caused about 1 in 10 deaths, but in countries where adult smoking is prevalent this proportion rises to 1 in 5 or more. Passive smoking is also a significant cause of death and disability among adults wherever smoking is prevalent, and although precise numbers are less certain than for active smoking, in Europe for example, deaths caused by passive smoking are equivalent to about 10% of the total attributable to active smoking. Passive smoking is also a major risk factor for death and disease in children, including fetal death, sudden infant death, lower respiratory infection, asthma, middle ear disease, and meningitis. Smoking also contributes to poverty and economic disadvantage as well as ill-health. Preventing smoking is therefore a major global public health priority.

VULNERABLE GROUPS

A number of vulnerable groups exist with regard to smoking prevalence which are a priority for tobacco control worldwide. Some have been discussed previously (for example those in lower socio-economic groups and in ethnic minority groups). Other vulnerable groups include those with mental health issues, prisoners, pregnant women, and adolescents.

Mental Health

The significance of tobacco smoking in those with mental illness is substantial. Issues related to the increased prevalence of smoking and severity of tobacco dependence, and the resulting impact on the lives of smokers with a mental illness have been confirmed in numerous international studies, notably in the United States, Australia, and New Zealand. Patients with mental illness are up to three times more likely to be smokers than the general population with smoking prevalence reaching over 70% for some severe disorders such as schizophrenia. Smokers with mental illness are likely to be heavier, more dependent smokers and have smoked for longer than smokers in the general population. Frequent and heavy smoking has detrimental impacts on the physical health of individuals with psychiatric conditions. It is responsible for a large proportion of the excess mortality experienced by those with mental health problems. With smoking being gradually denormalized in the general population, its current persistence and entrenchment in mental health settings contribute to a widening of tobacco-related health gaps in a highly vulnerable population. Those with mental health disorders are often as motivated to quit as the general population but face difficulties due to their high level of dependence and a lack of knowledge on the part of health professionals as to the link between smoking, mental illness, and treatments. Difficulties to address the 'smoking culture' in mental health settings such as those above and possible steps forward have recently been identified in national and international research.

Pregnancy

Smoking is the single largest modifiable risk factor for adverse outcomes in pregnancy and prenatal death. Women who smoke during pregnancy are more likely to be young, single, of low income and low educational attainment. Smoking has been linked to a number of serious health problems, including but not limited to ectopic pregnancy, increased risk of miscarriage, complications during labor, low birth weight, and stillbirth. In addition, an estimated 40% of all infant deaths have been estimated as being attributable to smoking. Babies born to mothers who smoke during pregnancy are more likely to suffer from poor intellectual development and physical growth and are more likely to have respiratory ailments. Smoking prevalence throughout pregnancy varies considerably around the world, but as examples figures of between 6 and 25% have been reported across Europe, 15–20% in both Australia and the United States and around 30% in South Africa. It is important to continue to encourage as many women as possible to quit smoking during pregnancy to minimize the health risks listed above.

Prisoners

A 1997 national survey of prisons in England and Wales revealed a smoking prevalence of 85% among male remand prisoners. Smoking plays a significant role in the prison setting, in particular as a relief from boredom and stress and for fostering a sense of group membership in an environment where smoking is very much the norm. These reasons also make cessation in this environment extremely difficult. Canada, Australia, and the United States are examples of countries which have introduced legislation against smoking in prisons which may contribute to reducing the prevalence, although the rulings have often come under challenge and in some cases have been overturned.

Adolescents

Adolescent smokers are of concern as they are the group who will have the largest impact on smoking prevalence in the future, and are also more likely to

drink alcohol or take illicit drugs in addition to tobacco use. Experimentation is an important predictor of future use, with the vast majority of regular smokers taking up the habit in adolescence. Reducing smoking among young people presents a challenge for health professionals as reasons for smoking are multi-faceted and largely driven by the home and social environment.

GENETICS

Despite decades of research to develop effective smoking-cessation treatments, a large proportion of smokers who attempt to quit fail, so that any improvement in the currently modest success rates achieved by smokers attempting to quit would have the potential to afford a considerable improvement in public health. There is substantial inter-individual variation in the level of nicotine dependence among cigarette smokers, as well as in therapeutic response to smoking-cessation pharmacotherapies. These observations have prompted a growing interest in genetic research to determine (1) whether individual differences in smoking behavior may be due to genetic influences and (2) whether the efficacy of different pharmacotherapies for the treatment of nicotine dependence is influenced by inherited variation in drug metabolizing enzymes and drug targets.

Genetic effects are known to influence smoking behavior, including cigarette consumption, but the identification of genetic variants robustly associated with smoking behavior phenotype has proved elusive. Recently, however, a number of genome-wide association (GWA) studies have provided robust evidence that single nucleotide polymorphism (SNP) variants with the *CHRNA5-A3-B4* gene cluster (which encodes the α5, α3, and β4 nicotinic acetylcholine receptor subunits) on chromosome 15q24 are associated with heaviness of smoking. Research in this area has reported an association with the rs16966968/rs1051730 SNPs (which are highly correlated and used interchangeably) with each additional copy of the risk allele associated with approximately one cigarette per day increased smoking quantity, corresponding to approximately 0.5% of phenotypic variance. Recently, a number of large collaborative efforts have replicated this association, and identified additional loci which appear to be associated with different aspects of smoking behavior, including smoking initiation, heaviness of smoking, and smoking cessation.

To date, only a limited number of pharmacogenetic studies of smoking-cessation pharmacotherapies have been published, principally investigating first-line medications such as NRT and sustained-release bupropion. A smaller number of studies investigated second-line medications such as venlafaxine. These pharmacogenetic analyses have typically focused on candidate genes in the dopamine, serotonin, and opioid pathways, or on genes coding for drug metabolizing enzymes (e.g. *CYP2A6* and *CYP2B6*).

It is clear that these findings are too preliminary to support the translational application of these findings to the clinical practice of smoking-cessation pharmacotherapy at the present time. In particular, there is a need for very large studies explicitly designed to examine pharmacogenetic pathways, implementing genome-wide analyses rather than focusing on single or a small number of variants. Newer medications (e.g. varenicline), as well as current second-line medications for smoking cessation (e.g. nortriptyline), will require investigation, and such studies should investigate associations with tolerability and side effects, given that there may exist sub-groups of individuals who respond well to certain medications which are typically not well tolerated.

The cost-effectiveness and acceptability of genetic testing, both among patients and clinicians, will also require investigation. Any additional costs generated by the consultation time required when providing feedback on the results of a genetic test may serve to render tailored intervention less cost-effective than the current model, despite the very low (and falling) costs of genotyping *per se*. There are several issues here, including assessing the clinical utility of such a treatment model, considering the preparedness of primary care physicians to implement such a model, determining the acceptability of such testing among patients, and ensuring that appropriate legal and regulatory frameworks are in place to ensure privacy and protect against genetic discrimination (e.g. by insurers).

Can we therefore know whether pharmacogenetics will ever help smokers quit? At present, the answer is "no", although there is some early promise. We will need to know a great deal more before we can begin to translate pharmacogenetic findings into practice, not the least of which is to determine which of those findings to date represent replicable, robust associations with treatment response. However, the pharmacogenetic investigation of smoking cessation has value beyond simply the potential for translation into clinical practice, and the implementation of tailored or personalized therapy. Such research can provide important insights into the mechanisms of nicotine addiction and smoking cessation, including response to pharmacotherapy.

TREATING TOBACCO DEPENDENCE

The idea that some smokers can become dependent on tobacco and may require treatment to stop smoking started to gain wider currency only in 1970s. The

awareness of health risks of smoking became widespread, a large number of smokers were quitting, and a special category of smoker became increasingly visible: someone who is highly motivated to stop but unable to do so.

A considerable number of psychological treatments were proposed, using mostly behavioral and some cognitive elements. A typical treatment program required smokers to monitor their smoking behavior, identify the most important cigarettes, and replace smoking with alternative coping strategies and behaviors. Despite the intuitive validity of this approach, little evidence emerged suggesting that any of the specific behavioral interventions are actually helpful in facilitating long-term abstinence. One exception to this was a method of aversive or 'rapid' smoking. This requires smokers to take frequent puffs on a series of cigarettes (typically one puff every 10 s) until smoking becomes aversive. The method seems to have an active ingredient and modest efficacy. It was superseded by NRTs in 1980s.

There are several NRT products currently available, including nicotine chewing gum, transdermal patch, nasal spray, lozenge, 'inhalator', and sublingual tablet ('microtab'). The first three of these were innovative when they first appeared, with the gum pioneering the new approach to treatment, the patch providing a user-friendly option and a potential for achieving higher nicotine concentrations, and the spray increasing the speed of nicotine delivery. Since 1980s, however, the field has been stagnant, with new products simply mimicking the delivery profile of the tried-and-tested gum.

The existing NRTs when combined with multi-session support almost double the success rates of the behavioral treatments, but overall their efficacy remains limited, reaching some 15% 1-year abstinence rates in studies with strict outcome criteria. Note that in this field, the choice of outcome definitions (e.g. validated or not, point prevalence versus sustained abstinence, etc.) has a stronger influence on reported 'success rates' than the methods used.

The anti-depressant bupropion was discovered to help smokers quit in the 1990s. Its efficacy is considered overall similar to NRT, but it seems to have a striking effect in some smokers. Compared to NRT, which has virtually no contraindications and no side effects other than local irritation, bupropion has contraindications and can cause insomnia and seizures. It remains popular in United States but in Europe, where UK tabloids ran a scare campaign against it, it is now used only rarely.

The selective α4β2 nicotinic acetylcholine receptor partial agonist varenicline has been available since 2006. It is more effective than bupropion and seems to be also more effective than NRT, at least in the short term. About a third of users experience nausea, but this is usually mild and does not deter smokers from carrying on. Media reports of erratic behavior and suicide in smokers using the drug led to warnings in the labeling, but data available so far suggest these were coincidences of the type to be expected in any medication used by millions of people.

All three medications show efficacy largely only when accompanied by behavioral support. Key elements of this appear to be a supervision of the actual quit date, regular monitoring of abstinence, ensuring patients understand what to expect from the medication and how to use it to its best advantage, reassurance regarding withdrawal discomfort, and maintaining patients' motivation to remain abstinent. In the United Kingdom and in an increasing number of other countries, the current prevailing approach to treating smokers is the Withdrawal Oriented Treatment (WOT), which combines pharmacological aids and motivational support.

CLINICAL TRIALS

There are two types of clinical trials: comparative and non-comparative. In the early phases of testing a pharmacotherapy for addiction, researchers use non-comparative trials to examine the effect of different doses on psychological and physical indices. Such trials can suggest a medication might be effective, but cannot provide definitive evidence. For this we need a comparative clinical trial where two or more treatments are given to participants and the frequency of occurrence of the outcome compared.

Although pharmacotherapy trials are relatively straightforward to describe, lots of trials in tobacco addiction concern psychological, behavioral, or societal interventions. These are often described as complex interventions. This means that the whole is more than the sum of the parts or that the various components depend on each other for their efficacy. Testing complex interventions follows a process similar to testing pharmacotherapy, described in the UK Medical Research Council Framework for Complex Interventions.

The huge variety of types of interventions and the phases of development make it impossible for all trials to be run the same way. However, the final test of any intervention is whether it can prevent people taking up tobacco or help people stop (or sometimes reduce) their use of tobacco and thus overcome their addiction. For this reason, the main academic society in the field, the Society for Research on Nicotine and Tobacco, has proposed guidelines to help researchers show that interventions work. For example, addiction is a relapsing disorder and the large majority of people who are

abstinent at one point in a treatment trial for addicted smokers will relapse. Thus the Society proposes that trials should show that participants have been abstinent for at least 6 months. This long-term criterion is much less important in prevention trials where, except in special populations, the majority will not take up smoking in the near future.

An important aspect of the analysis of clinical trials in addiction or in other fields is the intention to treat principle. Randomization of a sufficient number of participants ensures reasonable balance of both the known and unknown predictors of outcome (confounders). All other things being equal, both groups have an equal chance of the outcome occurring. In a treatment trial, participants whose attempt to stop smoking is failing and who abandon the program are unlikely to attend a clinic and are less likely to respond to attempts by the researcher to follow them up. Thus assessing the outcome in only those that attend follow-up is biased. We have empirical data from a typical treatment trial that smokers who decline follow-up are smoking. Thus, to cope with the missing data, we impute that people not followed up are smoking and there are detailed coding processes called the Russell Standard. This process does not work so well in some contexts: for example, people who are quitting using an Internet site are much less likely to be followed up but the assumption that they are all smoking is probably wrong. On the other hand, the assumption that those who are followed up have the same smoking experience as those who decline to be followed is also wrong. This issue is also not clear in prevention trials, where most children who are not followed are probably not smoking but data indicate that they are more likely to smoke than children who are followed up.

Another issue in tobacco addiction trials is that we can and usually do seek to confirm smoking status by biological tests. It is cheap and easy to measure exhaled carbon monoxide in smokers or do other tests on blood, urine, or saliva which would indicate active smoking. In a typical treatment trial, participants feel a strong pressure to maintain abstinence and when they fail to achieve it, feel pressure to claim this. Even more important, this sense of pressure can vary between the arms of the trial because, in behavioral trials, blinding is impossible, and, in trials of pharmacotherapy, participants can usually feel the effect of the medication so blinding is not completely successful. Consequently biological confirmation is helpful. Again, in trials where participants have little or no relationship with a therapist, the pressure to lie is lower and the feeling of wanting to prove abstinence makes biological confirmation harder to achieve and less necessary.

There are a large number of trials in prevention and treatment of tobacco addiction and the results of many of these trials are summarized in the Cochrane reviews by the Tobacco Addiction Review Group. This gives the best evidence of efficacy of both clinical and public health interventions in prevention and treatment of tobacco addiction.

TOBACCO POLICY AND CONTROL

In order to '*protect current and future generations from the devastating health, social, environmental and economic consequences of tobacco consumption and exposure to tobacco smoke reduce tobacco use internationally*' and in recognition of the globalization of the tobacco epidemic, the World Health Organization developed the world's first public health treaty, the Framework Convention on Tobacco Control (FCTC). Adopted by the World Health Assembly in 2003, it entered into force in February 2005 and since that time, 171 Parties (countries) have demonstrated political commitment not to undermine it, making it one of the most popular treaties in the history of the United Nations. The FCTC sets out a comprehensive strategy for reducing tobacco use and its harm. In its preamble it states that it represents a '*paradigm shift in developing a regulatory strategy to address addictive substances; in contrast to previous drug control treaties, the WHO FCTC asserts the importance of demand reduction strategies as well as supply issues*'.

There is a substantial evidence base that the optimum way to reduce tobacco use and disease in populations is through implementing a comprehensive tobacco control strategy involving a variety of different policies. The implementation of these evidence-based policies will have important health benefits which will save money for health services and most are relatively inexpensive and easy to implement. Nevertheless, the tobacco industry and its front groups routinely challenge tobacco control policies in the hope that this will deter other countries from adopting them.

A comprehensive strategy needs to be aimed at reducing uptake, increasing cessation of tobacco use and reducing the harmfulness of tobacco use to others and to those users who cannot or will not stop. The World Bank emphasized the importance of fiscal policies to increase the price of tobacco. It also demonstrated that encouraging existing tobacco users to stop is a better strategy for reducing the mortality caused by smoking over a relatively short time frame (the next few decades), rather than concentrating on preventing uptake among new users. In order to help as many tobacco users in a population to quit, policies and strategies with wide reach are required (such as price rises and mass media campaigns or media advocacy) as well as more intensive individual strategies, such as smoking-cessation clinics, which will reach fewer people but have a much greater

efficacy. Harm reduction policies have the least consensus, apart from measures to protect people from the risks of breathing other people's smoke. To achieve the latter, prohibiting smoking in work and public places has been shown to reduce exposure to the harmful effects of cigarette smoke and have an immediate impact on hospital admissions for heart attacks within a year of coming into effect.

Demand reduction strategies in the FCTC include: taxation policies which increase the price of tobacco over and above inflation on a regular basis; protection from exposure to tobacco smoke (such as smoke free public places); packaging and labeling of tobacco products (large pictorial health warnings are shown to be the most impactful); widespread education and public awareness; prohibition of tobacco advertizing, promotion, and sponsorship; regulation of the contents of tobacco products and of tobacco product disclosures; and the provision of tobacco dependence interventions. Supply reduction provisions include curtailing the illicit trade in tobacco products, reducing sales to and by minors and provision of support for economically viable alternative activities to tobacco growing and manufacture. Finally, the FCTC also includes a provision to address liability issues.

Monitoring tobacco policy implementation is important and a tobacco control scale has been developed to measure country progress with implementation of FCTC tobacco control policies. Article 14 of the FCTC states that countries who are signatories to the treaty should develop evidence-based treatment guidelines and take effective measures to promote adequate treatment for tobacco dependence. To date, however, relatively few countries have introduced smoking treatment programmes as part of their health care systems. One exception is the United Kingdom where a national network of smoking-cessation services now exists.

Given the widespread availability of cessation services, the United Kingdom provides an interesting example of how tobacco polices have been developed and implemented. In the United Kingdom, smoking-cessation services were developed on the basis of guidance originally published as a special edition of the journal *Thorax*. This guidance reviewed the effectiveness and cost-effectiveness of interventions to treat tobacco dependence and recommended that specialist clinics be established offering a combination of NRT (at the time, the only available stop smoking medication) and behavioral support. In 1998, the UK government published a white paper on tobacco entitled *Smoking Kills*. This proposed the development of a national network of cessation services delivered by the UK National Health Service (NHS). These services were initially piloted in areas of deprivation in England from 1999 and rolled out across the United Kingdom from 2000. The *Thorax* recommendations provided the basis for subsequent government guidance for the services, originally published in 2000. The services provide a combination of behavioral support from a trained adviser, delivered either one to one or in groups in a range of health service and community venues for between 6 and 8 weeks on average. Behavioral support is combined with access to a range of NRT products and/or bupropion or varenicline. All medications are available on prescription at little or no cost to the smoker.

Since their establishment around 4.7 million smokers have set a quit date with the services in England. Of these, 2.5 million have reported that they stopped smoking in the short term (4-week post quit date) with 1.6 million of these successful quit attempts confirmed by biochemical validation (CO monitoring). Research has examined the longer-term effectiveness of the services, finding a validated 1-year quit rate of 14.6%. This equates to a fourfold increase in the chances of quitting compared with trying to stop unaided. This is equivalent to around 680 000 smokers stopping in the longer term amongst all those who have accessed the services in England since their establishment.

Research has also demonstrated that smoking-cessation services can be effective in reducing inequalities in health. The evidence base to guide the development of interventions to reduce inequalities is extremely limited, but treatment for tobacco dependence provides one of the only clear examples of what works. Studies have demonstrated that smoking-cessation services are effective in reversing the inverse care law (that states that health care is more readily available to affluent than deprived groups). They do this by reaching and treating proportionally more smokers in disadvantaged areas than in more affluent areas. One study examined quit rates amongst 1.5 million NHS stop smoking service clients that had used the services between 2003 and 2006. This found that although smokers from poorer areas were less likely to be successful in quitting, the services were treating a far higher proportion of these smokers. This resulted in a reduction in absolute and relative gaps in smoking rates between disadvantaged and more affluent areas – thereby making a contribution to reducing inequalities in health caused by smoking.

To date, smoking-cessation services have largely been established in developed countries such as the United Kingdom where smoking rates may be higher and where the health care system is well developed. In developing nations, there is some evidence that the integration of tobacco dependence treatment into existing programs such maternity services or clinics to treat tuberculosis, or HIV/AIDS, for example, may be

a more effective way of reaching smokers than through specialist cessation services.

TOBACCO MARKETING

Why do so many young people still take up smoking? Why do so many continue with the habit for decades thereafter and, in many cases, until it kills them? And why do both phenomena persist despite widespread knowledge of the hazards involved and the best efforts of public health? Scientists cite dozens of psycho-social and biological factors, including demographics, peer pressure, parental modeling, economics, disadvantage, gender and, of course, addiction – all of which interact in a complex and damaging system that demands massive and multi-faceted effort to counteract it. However, a much simpler explanation is that smoking continues to thrive because a handful of large, multinational tobacco corporations use their marketing muscle to ensure that it does.

For many, the word marketing is synonymous with mass media advertising, and this can be a powerful tool. The Marlboro Cowboy and Joe Camel brand images are just the visible tip of a beguiling iceberg. Billboards and press ads are combined with a host of other less obvious communications, including sports and arts sponsorship, digital media, merchandising, point of sale display, and pack design. The latter two have become increasingly important in jurisdictions like the United Kingdom and Australia which have banned other sorts of advertising.

This mix of 'marketing communications' is nested within the rest of the marketing effort of product design, distribution, and pricing, with the combined aim of making an offering to smokers and potential smokers that is as attractive, ubiquitous, and as affordable as possible. So young and old are perpetually tempted to try, and continue consuming, an appealing array of beautifully packaged and displayed tobacco products. This effort is strategically and coherently deployed for many years (the Lucky Strike brand, for instance, is a 100 years old) and the physical offering is developed into a much broader and emotionally satisfying brand. Hence a bit of dried leaf in a paper tube becomes an evocative lifestyle choice and powerful statement of identity – and the alchemists' dream of turning base ingredients into gold is fulfilled.

But the tobacco marketers' work is not complete. They have read John Donne, understand that no man (or woman) is an island, and that social context has a big impact on individual choice. They therefore put just as much effort into influencing the behavior of stakeholders, policy makers, and power-brokers as they do into influencing the behavior of teenagers, but in this case, rather than starting to smoke, they want them to resist tobacco control measures such as marketing restrictions or smoke free ordinances; to treat their industry with respect and to encourage pro-tobacco social norms. Thus the tobacco industry recruits not just individuals but the system to its cause.

SUMMARY

Nicotine is considered to be the main addictive ingredient in tobacco. It reinforces smoking behavior by indirectly stimulating dopamine transmission in the reward pathways of the brain. Furthermore, chronic nicotine administration induces neurobiological changes that manifest in an aversive withdrawal syndrome when abstinent that motivates the individual to smoke in order to avoid a negative physiological and emotional state.

Smoking is the most popular vehicle of recreational nicotine consumption, due to its fast pharmacological action. However, smoking is associated with numerous negative health consequences, including cancers, cardiovascular disease, and respiratory illness. Increasing awareness of the risks, and heightened social stigmatization of smoking in recent years, may explain the steady decrease in smoking rates in developed countries, although rates are stagnating or increasing in developing countries. Within societies, smoking is particularly prevalent or problematic among certain groups. For example, individuals of low SES, adolescents, pregnant women, and mental health patients are particularly vulnerable.

Numerous psychological treatments have been proposed to treat nicotine dependence, using behavioral and cognitive techniques. More recently, pharmacotherapies have proved efficacious, particularly when combined with behavioral support. The most common are NRT, bupropion, and varenicline. Of these, varenicline is generally considered the most effective, although there is individual variability in the response to these drugs, and cessation rates remain low. In order to improve cessation rates, there has been growing interest in genetic research to further understand individual variations in smoking behavior and response to treatment, which could result in the introduction of targeted cessation treatments, although the cost-effectiveness of such interventions has been questioned.

In recognition of the need to reduce smoking rates and smoking-related harms, tobacco control strategies are being increasingly developed and enforced throughout the world. These include increasing the taxes, restricting smoking in public-areas, incorporating health warnings on packets, and increasing the availability of smoking-cessation services. Legislative restrictions have also been placed on the marketing of tobacco products in

order to reduce the allure of cigarettes to current smokers but also to non-smokers that may be tempted to smoke.

Acknowledgment

The authors are all members of the UK Centre for Tobacco Control Studies, a UKCRC Centre for Public Health Excellence. Funding from the British Heart Foundation, Cancer Research UK, the Economic and Social Research Council, the Medical Research Council, and the National Institute of Health Research, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged.

SEE ALSO

Gender Differences, Impact of Substance Use on the Course of Serious Mental Disorders, Epidemiology of Adolescent and Young Adult Alcohol, Tobacco, and Drug Use and Misuse in the United States

Glossary

Bupropion an atypical anti-depressant drug that is also a licensed smoking-cessation pharmacotherapy due to its ability to reduce the severity of smoking craving and withdrawal symptoms.

Chronic obstructive pulmonary disease (COPD) a life-threatening respiratory disease characterized by chronic obstruction of lung airflow.

Cytochrome P450 2A6 (CYP2A6) one of the Cytochrome P450 enzymes. Involved in the metabolism of nicotine to its primary metabolite cotinine.

Cytochrome P450 2B6 (CYP2B6) one of the Cytochrome P450 enzymes. Involved in the metabolism of nicotine to its primary metabolite cotinine.

Dopamine one of the catecholamine neurotransmitters in the brain. Plays an important role in the reward systems of the brain.

Genome-wide association (GWA) studies methodological approach employed in genetic epidemiology. Involves analysis of genetic markers across the whole genome (i.e. full set of chromosomes). Used to identify associations between genetic variations (i.e. genotypes) and specific phenotypes of interest (e.g. disease, personality traits).

Monoamine oxidase inhibitors (MAOIs) a class of chemicals that increase the levels of monoamine neurotransmitters by inhibiting the enzyme that metabolizes them (monoamine oxidase). Used clinically to treat depression.

Nicotine replacement therapy (NRT) family of smoking-cessation pharmacotherapies that involve administration of nicotine by non-smoking means. Common varieties include: transdermal patch, lozenge and gum.

Nortriptyline an anti-depressant drug that has also been used to aid smoking cessation. Generally less prescribed than other pharmacotherapies for smoking as its use is commonly associated with side effects.

Nucleus accumbens (NAcc) brain region within the forebrain that is part of the mesolimbic dopamine pathway and plays an important role in motivation and reward.

Point of sale the physical location at which goods are bought/sold, such as a shop checkout.

Point prevalence the number of cases within a clinically relevant category (e.g. smoking abstinent) when measured at a specific point in time.

Single nucleotide polymorphism (SNP) genetic variation within a DNA sequence that occurs when a single nucleotide in a genome differs between individuals or between paired chromosomes in the same individual.

Varenicline a nicotine receptor partial agonist used as a smoking-cessation pharmacotherapy. Acts by reducing nicotine withdrawal and the pleasurable effects of smoking.

Venlafaxine an anti-depressant and anti-anxiety drug that has shown to be beneficial as a smoking-cessation aid, particularly when paired with other pharmacotherapies such as nicotine replacement therapy.

Ventral tegmental area (VTA) sub-cortical brain region where the dopaminergic cell bodies of the mesolimbic reward pathway are located.

Withdrawal Oriented Treatment (WOT) a therapeutic approach to smoking cessation that focuses on helping nicotine-dependent individuals overcome the negative consequences of nicotine deprivation.

Further Reading

Bauld, L., Bell, K., McCullough, L., Richardson, L., Greaves, L., 2009. The effectiveness of NHS smoking cessation services: a systematic review. Journal of Public Health 32, 71–82.

Hughes, J.R., 2003. Motivating and helping smokers to stop smoking. Journal of General Internal Medicine 18, 1053–1057.

Jha, P., Chaloupka, F.J. (Eds.), 2000. Tobacco Control in Developing Countries. Oxford University Press, Oxford.

McNeill, A., Raw, M., Whybrow, J., Bailey, P., 2005. A national strategy for smoking cessation treatment in England. Addiction 100 (Suppl. 2), 1–11.

Munafò, M.R., Shields, A.E., Berrettini, W.H., Patterson, F., Lerman, C., 2005. Pharmacogenetics and nicotine addiction treatment. Pharmacogenomics 6, 211–223.

Raw, M., McNeill, A., West, R., 1998. Smoking cessation guidelines and their cost-effectiveness. Thorax 53 (Suppl. 5), S1–S19.

Shafey, O., Dolwick, S., Guindon, G., 2003. Tobacco Control Country Profile 2003. American Cancer Society, Atlanta, GA.

Shafey, O., Eriksen, M., Ross, H., Mackay, J., 2009. The Tobacco Atlas, third ed.). American Cancer Society, Atlanta.

Shahab, L., 2008. The epidemiology of smoking: a growing concern. In: Miravitlles, M. (Ed.), Hot Topics in Respiratory Medicine (Issue 8). FB Communications, Modena, pp. 7–14.

Thorgeirsson, T.E., Geller, F., Sulem, P., et al., 2008. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature 452, 638–642.

West, R., 2009. The multiple facets of cigarette addiction and what they mean for encouraging and helping smokers to stop. COPD: The Journal of Chronic Obstructive Disease 6, 277–283.

Relevant Websites

Cochrane Clinical Trials Library – www.thecochranelibrary.com.
The Information Centre for Health and Social Care – www.ic.nhs.uk/pubs/sss0910.
Medical Research Council – www.mrc.ac.uk.
Society for Research on Nicotine and Tobacco – www.srnt.org.
WHO Framework on Tobacco Control – www.who.int/fctc.
World Health Organization – www.who.int.
UK Department of Health – www.dh.gov.uk.

CHAPTER

78

Caffeine and Caffeinated Energy Drinks

Andrew Paul Smith

Centre for Occupational and Health Psychology, School of Psychology, Cardiff University, Cardiff, UK

OUTLINE

BACKGROUND

The effects of caffeine have been widely studied and a number of recent reviews cover different outcome measures. Indeed, the present author has written a number of reviews of the behavioral effects of caffeine and has suggested that a cost–benefit analysis is an appropriate way of organizing the literature. The first part of this article will summarize research which demonstrates beneficial effects of caffeine. Earlier reviews suggested that the behavioral effects of caffeine are often positive except when one considers very large doses and sensitive individuals. The issue of dependence and negative effects associated with withdrawal clearly represent an area where one needs to assess possible costs of caffeine consumption. Consumption of energy drinks by children has become an important recent issue. These drinks often have high levels of caffeine and are being consumed by a potentially vulnerable sample. An overview of our current knowledge of these two topics will be given in the second part of the article.

Considerable detail about sources of caffeine, the pharmacology of caffeine, and levels of consumption is available elsewhere. The next section provides a very brief summary of these topics. Caffeine (1,3,7-trimethylxanthine) is a member of a class of naturally occurring substances termed methylxanthines. Plasma levels of caffeine peak 15–45 min after ingestion and the half-life is between 5 and 6 h. A variety of factors influence the metabolism. For example, in pregnant women the half-life can increase to 18 h. Oral contraceptive use

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00078-4

increases the half-life to 11 h whereas in cigarette smokers the half-life is only about 3 h which may account for the high level of consumption found in this group.

Caffeine acts by blocking the effects of the naturally occurring neuromodulator adenosine. This produces an increase in central nervous system (CNS) activity which is associated with changes in many neurotransmitter systems. Caffeine has other CNS effects (e.g. influencing blood flow to the brain). Many of these CNS effects of caffeine are unlikely to occur with the amounts consumed by humans.

Caffeine occurs naturally in a number of foods, is employed as a food additive, and is added to medications. It occurs naturally in coffee, tea, and cocoa and the exact amount present will depend on growing conditions and preparation. Some rough approximations of the caffeine content of products are: filter coffee – 100–150 mg (5 oz cup); instant coffee – 50–60 mg; tea – 35–45 mg; milk chocolate – up to 15 mg; dark chocolate – up to 35 mg. It is also added to soft drinks (e.g. Cola 40 mg in a 12 oz serving). Higher amounts are added to energy drinks (70 mg+). Caffeine is also added to over-the-counter (OTC) medications (e.g. to analgesics, usually in the region of 30–50 mg per tablet). Caffeine tablets can also be purchased and the recommended dose to increase alertness is 100–200 mg. Several estimates suggest that average intake is about 200 mg per day. Problems usually only occur when excessive amounts are consumed (500 mg a day plus) and these problems usually take the form of an increase in anxiety. Some individuals are very sensitive to effects of caffeine and even small amounts can cause adverse reactions. Most individuals control their consumption of caffeine. For example, consumption usually occurs when alertness is reduced (e.g. early in the morning; after prolonged work; after lunch) and is reduced at times when high alertness is undesirable (e.g. before going to sleep).

BENEFICIAL EFFECTS OF CAFFEINE

In a recent review of the literature on possible beneficial effects of human caffeine consumption, Michael Glade concludes that moderate amounts of caffeine lead to the following benefits:

1. Increased energy availability.
2. Increased daily energy expenditure.
3. Decreased fatigue.
4. Decreased sense of effort associated with physical activity.
5. Enhanced physical performance.
6. Enhanced motor performance.
7. Enhanced cognitive performance.
8. Increased alertness, wakefulness, and feelings of energy.
9. Decreased mental fatigue.
10. Faster reactions.
11. Increased accuracy of reactions.
12. Increased ability to concentrate and focus attention.
13. Enhanced short-term memory.
14. Increased ability to solve problems requiring reasoning.
15. Increased ability to make the correct decisions.
16. Enhanced cognitive functioning capabilities and neuromuscular coordination.

CAFFEINE AND SPORTS PERFORMANCE

A recent position paper on caffeine and sports performance from the International Society of Sports Nutrition can be summarized as follows:

1. Caffeine is effective for enhancing sport performance in trained athletes when consumed in low-to-moderate dosages (~3–6 mg kg^{-1}) and overall does not result in further enhancement in performance when consumed in higher dosages ($\geq$9 mg kg^{-1}).
2. Caffeine exerts a greater ergogenic effect when consumed in an anhydrous state as compared to coffee.
3. It has been shown that caffeine can enhance vigilance during bouts of extended exhaustive exercise, as well as periods of sustained sleep deprivation.
4. Caffeine is ergogenic for sustained maximal endurance exercise, and has been shown to be highly effective for time-trial performance.
5. Caffeine supplementation is beneficial for high-intensity exercise, including team sports such as soccer and rugby, both of which are categorized by intermittent activity within a period of prolonged duration.
6. The literature is equivocal when considering the effects of caffeine supplementation on strength-power performance, and additional research in this area is warranted.
7. The scientific literature does not support caffeine-induced diuresis during exercise, or any harmful change in fluid balance that would negatively affect performance.

CAFFEINE AND SLEEP DEPRIVATION

The American Academy of Sleep Medicine has examined the efficacy and safety of caffeine use during sleep loss and their main conclusions are summarized below:

The most commonly reported measure used to study this topic is the ability to stay awake or fall asleep. Fourteen out of 15 studies reviewed have shown increased wakefulness measured by sleep latency tests following ingestion of caffeine by sleep-deprived volunteers.

Choice reaction time performance of sleep-deprived individuals has been improved by caffeine in eight studies. Similarly, working memory performance of sleep-deprived individuals has been shown to be improved by caffeine in over 10 studies. The effects of caffeine during sleep loss have been examined over a dose range from 75 to 1200 mg per 24 h. Recommended doses are usually in the range of 200–300 mg, mainly because side effects are more prevalent with higher doses. Caffeine administration typically improves performance during sleep loss as compared with placebo, but performance and alertness often continue to decline even when caffeine is given due to further sleep loss, circadian rhythms, and caffeine half-life.

Subjective alertness decreases with sleep loss and ratings of fatigue increase. Studies that have monitored mood typically show that caffeine ameliorates these subjective changes with a similar time course to that seen for performance variables. However, some studies have not obtained these results and it has been suggested that some "subjective tolerance" to caffeine may develop with prolonged testing.

Studies have generally shown that doses of 200–300 mg caffeine produce few side effects whereas higher doses (600 mg+) may increase mild symptoms (e.g. gastrointestinal upset, nervousness, muscle twitching). Based on these findings, the review concluded that caffeine can increase alertness and improve performance at doses of 75–150 mg after acute restriction of sleep and at doses of 200–600 after a night or more of total sleep loss. Caffeine is unlikely to have major disruptive effects on sleep that follows 8 h or longer after administration. Prolonged administration is not recommended due to the increasing likelihood of side effects with high doses.

CAFFEINE AND SUSTAINED MILITARY OPERATIONS

A number of studies have examined effects of caffeine in sustained military operations. Harris Lieberman and colleagues conclude that "When cognitive performance is critical and must be maintained during exposure to severe stress, administration of caffeine may provide a significant advantage".

Other research has examined beneficial effects of caffeinated tube food on pilot performance during a 9-h simulated U-2 mission. The results showed that the caffeinated tube food (200 mg caffeine consumed every 4 h) maintained cognitive performance at baseline levels over a 9-h overnight period. Research has considered both cognitive and physical performance measures in sustained operations (e.g. performance during 27 h of sustained wakefulness in Special Forces personnel). The results showed that caffeine (200 mg caffeinated gum administered on three occasions) maintained performance of a reconnaissance vigilance task and also improved running times compared to placebo. A similar study was conducted over a period of 4 days and three nights of sustained operations. The results showed caffeine maintained both vigilance and physical performance during sustained operations that require periods of overnight wakefulness and restricted opportunities for daytime sleep.

CAFFEINE AND SHIFT WORK

A recent review has considered the effects of caffeine for preventing injuries, errors, and cognitive problems caused by impaired alertness in persons doing shift work. Thirteen trials were included but none measured injury, two measured error, and the remaining trials assessed cognitive performance. The trials assessing the impact on errors found that caffeine significantly reduced the number of errors. Caffeine improved concept formation and reasoning, memory, orientation, attention, and perception. The results were largely from studies involving young participants under simulated conditions and further research is needed on older workers and real world shift work. The authors conclude that "Based on the current evidence, there is no reason for healthy individuals who already use caffeine within recommended levels to improve their alertness to stop doing so."

CAFFEINE AND SAFETY AT WORK

Research by the author has examined the impact of habitual caffeine consumption on performance and safety at work. In the first study volunteers, all of whom were regular caffeine consumers, rated their alertness and carried out a simple reaction time task before and after work on a Monday and Friday. Caffeine consumption during the day was recorded and volunteers were sub-divided into low and high consumers on the basis of a median split (220 mg day^{-1}). The results showed that those who consumed higher levels of caffeine reported significantly greater increases in alertness over the working day and a significantly smaller slowing of reaction time.

The second study involved secondary analyses of a database formed by combining the Bristol Stress and Health at Work and Cardiff Health and Safety at Work studies. In the first analyses associations between caffeine consumption and frequency of cognitive failures were examined in a sample of 1253 white-collar workers. The second set of analyses examined associations between caffeine consumption and accidents at work in a sample of 1555 workers who were especially at risk of having an accident. The results from the second study demonstrated significant associations between caffeine consumption and fewer cognitive failures and accidents at work. After controlling for possible confounding factors, it was found that higher caffeine consumption was associated with about half the risk of frequent/very frequent cognitive failures and a similar reduction in risk for accidents at work. Overall, the results from the three analyses confirmed that caffeine consumption may have benefits for performance and safety at work.

CAFFEINE AND HUMAN ERROR

Other research by the author has involved secondary analyses of a large epidemiological database to examine associations between caffeine consumption and cognitive failures (errors of memory, attention, and action) in a non-working sample. Associations between caffeine consumption and physical and mental health problems were also examined. After controlling for possible confounding factors significant associations between caffeine consumption and fewer cognitive failures were observed. Overall, the results show that caffeine consumption may benefit cognitive functioning in a non-working population. This confirms earlier findings from working samples. This beneficial effect of caffeine was not associated with negative health consequences.

CAFFEINE AND DRIVING

A number of studies by Jim Horne and colleagues have examined (1) the efficacy of 200 mg caffeine with restricted and completely sleep-deprived drivers, (2) an energy drink containing caffeine with sleep-restricted drivers, and (3) "a functional energy drink" and sleep-restricted drivers. The results from these studies showed that caffeine generally reduced the impaired driving performance that was seen in sleepy drivers given placebo.

Others have extended these results by examining the effects of sleepiness and caffeine on real-life driving. Extended driving and sleepiness resulted in an increase in lane crossing, which was reduced by 200 mg caffeine. As shown in previous sections of this chapter, fatigue can be induced in a number of ways. One study involved 1 h of simulated driving before and after either caffeine or placebo. In addition, volunteers carried out a battery of tasks measuring subjective alertness and sustained attention. Caffeine reduced steering variability (which in real-life driving may lead to lane crossing) and led to an increase in subjective alertness and improved cognitive vigilance. This suggests that results found after caffeine with artificial laboratory tasks may be applicable to real-life activities involving similar functions.

Driving performance can be impaired by a number of factors, the most widely studied being alcohol. One study examined whether caffeine would reduce an alcohol-induced impairment of simulated driving. The results suggested that caffeine may increase alertness and improve reaction time after alcohol use but will not completely counteract the alcohol impairments seen in driving.

Jack James and colleagues argue that many of the effects of caffeine seen in studies of driving can be interpreted in terms of reversal of the effects of caffeine withdrawal. One method of distinguishing a benefit of caffeine from a reversal of caffeine withdrawal is to compare caffeine consumers with non-consumers. There is a need to do this using an epidemiological approach to examine associations between caffeine consumption and road traffic accidents.

Research by the author has examined a community sample from South Wales ($N = 6648$). These respondents provided information on involvement in road traffic accidents. 3.6% of non-consumers of caffeine were involved in a road accident requiring medical attention compared to only 2.2% of caffeine consumers. Logistic regressions, including demographic, lifestyle, and psychosocial characteristics showed that consumption of caffeine nearly halved the risk of being in a road accident (odds ratio [OR] = 0.58 confidence intervals [CI]: 0.35, 0.98). This result confirms previous research showing that caffeine reduces the risk of accidents (at work) and supports the existing literature and information campaigns about the positive benefits of caffeine for road safety.

CAFFEINE AND DEMENTIA

Several epidemiological studies have examined associations between consumption of caffeine and dementia. A recent systematic review and meta-analysis considered nine cohort and two case control studies. The outcomes examined were Alzheimer's disease (four studies), dementia or cognitive impairment (two studies), and cognitive decline (three studies). The

summary relative risk for the association between caffeine intake and the different cognitive measures was 0.84 [95% CI: 0.72–0.99]. This suggests a trend toward a protective effect of caffeine but the large methodological heterogeneity across a small number of studies precludes more definitive conclusions.

CAFFEINE AND ANXIETY AND DEPRESSION

Caffeinism refers to a constellation of symptoms associated with very high caffeine intake that are virtually indistinguishable from severe chronic anxiety. Caffeinism is usually associated with daily intakes of between 1000 and 1500 mg. However, it appears to be a rather specific condition and there is little evidence for correlations between caffeine intake and anxiety in either nonclinical volunteers or psychiatric outpatients. Other research has investigated whether caffeine is capable of increasing the anxiety induced by other stressors. It has been found that 400 mg of caffeine increased anxiety when paired with a stressful task. However, other research has not been able to provide any evidence of interactive effects of caffeine and stress.

Recent research has shown an association between ADORA2A and DRD2 polymorphisms and caffeine-induced anxiety. Adenosine receptors functionally interact with dopamine receptors in the brain. Functional polymorphisms in the genes for either adenosine or dopamine receptors may, therefore, affect responses to caffeine. A recent study found that 50 mg caffeine did not increase anxiety in any individuals whereas 450 mg caffeine increased it in the majority of the volunteers. With a dose of 150 mg caffeine anxiety was associated with ADORA2A and DRD2 polymorphisms.

In contrast, moderate caffeine intake has been associated with fewer depressive symptoms and a lower risk of suicide. This effect of caffeine on depression may have other knock on effects with regards to health. The author conducted secondary analyses of a large epidemiological database ($N = 2750$) to examine associations between caffeine and both chronic and acute health outcomes. Many of the initial associations between caffeine and health were no longer significant when potential confounders were examined. However, caffeine consumption was still significantly associated with reduced depression in the final regressions. Caffeine consumption was also associated in a dose response fashion with fewer upper respiratory tract symptoms. This suggests that caffeine may influence the immune system, either directly, or by reducing depression (a well-established risk factor for immunosuppression). Other research by the author has shown that caffeine removes the malaise (fatigue, psychomotor slowing) associated with minor illnesses such as the common cold.

CHRONIC HEALTH EFFECTS OF CAFFEINE CONSUMPTION

It is important to conduct a cost–benefit analysis when considering the effects of caffeine. Benefits usually refer to behavioral outcomes and costs reflect possible long-term health effects. Caffeine has been linked with a range of possible health problems but most of these associations are not significant when confounding factors are adjusted for. Indeed, in recent years the trend has been for suggestions that caffeine may have health benefits. It has been suggested that studies of caffeine have played a key part in defining the role of adenosine receptors, phosphodiesterases, and calcium release channels in physiological processes. Caffeine and various analogs, the latter designed to enhance potency and selectivity toward specific biological targets, are potential therapeutic agents for intervention in Alzheimer's disease, asthma, cancer, diabetes, and Parkinson's disease.

CAFFEINE: AN ATYPICAL DRUG OF DEPENDENCE

Roland Griffiths and colleagues have suggested that caffeine is an excellent model compound for understanding drugs of abuse/dependence. Caffeine can be shown to act as a reinforcing agent, a criterion for dependence, under certain conditions. However, the level of responding is lower than that maintained by addictive drugs such as cocaine and amphetamine and there is little or no evidence for upward dose adjustment. Indeed, it is unclear about the contribution of desirable stimulatory effects and undesirable withdrawal symptoms in the reinforcing properties of caffeine.

It has been suggested that caffeine shares four behavioral pharmacological effects with classic drugs of abuse/dependence: reinforcing effects, discriminative/subjective effects, tolerance, and physical dependence. Similarly, it has been suggested that there are some people who report a compulsive pattern of caffeine use and are physiologically dependent on caffeine and there is evidence that just under 20% of caffeine users show some degree of dependence although this is small compared with nicotine (90% of users show dependence) and at a similar level to alcohol (where 14% show a lifetime prevalence of dependence). The most important issue is the severity of the harmful consequences associated with dependence. Compared to nicotine or alcohol the risks associated with moderate

caffeine consumption are generally low. Although caffeine fulfills some of the criteria for drug dependence and shares with amphetamines and cocaine a certain specificity of action on the cerebral dopaminergic system, the methylxanthine does not act on the dopaminergic structures related to reward, motivation, and addiction.

CAFFEINE WITHDRAWAL

Caffeine withdrawal is typically associated with symptoms of headache and drowsiness. These symptoms generally begin slowly, maximize after 1–2 days and are over within a few days. Many of the early studies of this topic used small samples and if one adjusted for the number of analyses carried out one would find few significant effects. The studies also have other undesirable features (e.g. in one study – the subjects were also the authors of the paper). The frequency of caffeine withdrawal has been examined in a population-based survey and in a controlled, blinded experiment. In the survey of over 11 000 people, 61% reported daily caffeine consumption and 11% of the caffeine consumers reported symptoms upon stopping caffeine. When volunteers were unaware that the focus of the study was caffeine withdrawal, reports of symptoms associated with withdrawal were less frequent. Indeed, in another double-blind study by the present author, caffeine withdrawal was associated with an increase in reports of headache but those who continued to consume caffeine also reported more headaches as the study progressed. Volunteers in this study were not very good at discriminating whether they were in the caffeine or no caffeine groups and this suggests that symptoms of caffeine withdrawal may only be apparent if volunteers know that caffeine has been withdrawn. Such a result was obtained in further research by the author which suggests that the dependence associated with caffeine may largely reflect the knowledge that caffeine has been withdrawn rather than a pharmacological dependence.

BENEFICIAL EFFECTS OF CAFFEINE OR REMOVAL OF NEGATIVE EFFECTS OF WITHDRAWAL?

Overall, the previous sections confirm that the effects of caffeine on performance are largely beneficial. However, this view has been questioned by Jack James who argues that the beneficial effects of caffeine are really only removal of negative effects produced by caffeine withdrawal. The author has argued against this general view of caffeine effects on a number of grounds. First, it cannot account for the behavioral effects seen in animals or non-consumers where withdrawal cannot occur. Second, caffeine withdrawal cannot account for behavioral changes following caffeine consumption after a short period of abstinence or the greater effects of caffeine when arousal is low. Finally, claims about the negative effects of caffeine withdrawal require closer examination as they can often be interpreted in ways other than caffeine dependence (e.g. expectancy). Indeed, in most of the studies that have demonstrated increases in negative affect following caffeine withdrawal, the volunteers have not been blind but have been told or even instructed to abstain from caffeine. This is clearly very different from the double-blind methodology typically used to study effects of caffeine challenge.

The view that beneficial effects of caffeine reflect degraded performance in the caffeine-free conditions crucially depends on the strength of the evidence for withdrawal effects. Jack James states that “there is an extensive literature showing that caffeine withdrawal has significant adverse effects on human performance”. If one examines the details of the studies cited to support this view one finds that some of them do not even examine performance, and that where they do, any effects are selective, not very pronounced, and largely unrelated to the beneficial effects of caffeine reported in the literature.

Peter Rogers and colleagues have reviewed a number of studies of caffeine withdrawal and performance. They conclude that “ in a review of recent studies we find no unequivocal evidence of impaired psychomotor performance associated with caffeine withdrawal”. Indeed, they found that caffeine improved performance in both deprived volunteers and non-consumers. Furthermore, other studies which suggest that withdrawal may impair performance can be interpreted in other ways than deprivation (e.g. changes in state).

The effects of caffeine withdrawal are still controversial. One study showed that caffeine withdrawal impaired short-term memory performance but caffeine ingestion had no effect. In contrast, research by the author has shown that caffeine improved attention in both those who had been deprived of caffeine for a short period and those who had no caffeine for 7 days. Other studies suggest that effects of withdrawal are restricted to mood and that performance is unaltered. Like many areas of caffeine research, some of the effects that have been attributed to withdrawal are open to other interpretations. For example, some studies have compared days when mid-morning coffee was either caffeinated or de-caffeinated. Caffeine consumption was associated with better performance and mood. The authors interpret this as a negative effect of caffeine withdrawal whereas one could equally interpret it as a positive effect

of caffeine. Other studies of caffeine withdrawal effects have methodological problems such as the lack of pre-drink baselines or failure to consider possible asymmetric transfer when using within subject designs.

Caffeine withdrawal has been widely studied because it is meant to provide crucial evidence on whether caffeine is addictive or leads to some kind of dependence. The most frequent outcome measure has been reporting of headache, but mood has been examined in other studies. Research has shown that caffeine deprivation led to increased reporting of stress by heavy coffee drinkers. This has been confirmed in another study which showed that caffeine withdrawal was associated with feelings of fatigue and decreased feelings of alertness. Indeed, results show that about 10% of volunteers with a moderate daily intake (235 mg day^{-1}) reported increased depression and anxiety when caffeine was withdrawn. Other research has examined the effects of varying time periods of caffeine deprivation (90 min, overnight and 7 days) on mood. The results showed that overnight caffeine deprivation produced dysphoric symptoms and these mood effects were reduced, but still present, after longer term abstinence. However, close examination of the results does not support this conclusion with only one of the 17 mood scales showing a significant effect.

Recent research in this area has been concerned with two main topics, namely what underlies the increase in symptoms following caffeine withdrawal, and, secondly, whether the positive effects of caffeine reflect removal of negative effects of withdrawal. Peter Dews and colleagues have considered factors underlying caffeine withdrawal and conclude that "non-pharmacological factors related to knowledge and expectation are the prime determinants of symptoms and their reported prevalence on withdrawal of caffeine after regular consumption".

In contrast, some researchers still suggest that caffeine only has beneficial effects on performance when the person has had caffeine withdrawn. One study reported that caffeine improved performance on a sustained attention task and increased rated alertness when volunteers had been caffeine deprived but had no such effects when they were no longer deprived. However, the results showed an effect of order of treatments with those who received caffeine first continuing to show better performance even when subsequently given placebo.

Research by the author has examined effects of caffeine in the evening after a day of normal caffeine consumption. Caffeine improved performance which casts doubt on the view that reversal of caffeine withdrawal is a major component underlying effects on performance. Further evidence against the caffeine withdrawal explanation comes from recent studies of non-consumers. These studies not only detected few negative effects of withdrawal but also showed that caffeine improved the performance of both withdrawn consumers and non-consumers, a finding that argues strongly against the withdrawal reversal explanation.

Other research has compared the effects of caffeine following abstinence and normal caffeine use. Caffeine had a greater effect on mood in the abstained state. The authors also suggest that choice reaction time showed a similar effect although this would not be significant if adjustments were made for multiple statistical tests. Other aspects of performance showed significant effects of caffeine in both abstained and normal caffeine consumption conditions. Failure to adjust for multiple testing is a common problem in this area of research. In addition, it is often unclear why specific sample sizes or tests are used. Consideration of these factors leads to a very different interpretation of some of the literature. For example, Heatherley and colleagues claim that cognitive performance is only improved by caffeine after 8 h of abstinence. Adjustment for multiple testing shows that none of the effects of caffeine are significant which reflects the low power of the study and failure to covary baseline data. Similarly, claims that 9 to 11-year-old children show negative symptoms of withdrawal which are reversed by caffeine do not hold up when adjustments are made for the number of statistical tests conducted.

A review of caffeine withdrawal has been conducted by Roland Griffiths with a view to validate specific symptoms and signs and to define important features of the syndrome. The review covered 57 experimental and 9 survey studies. Symptoms associated with caffeine withdrawal were: headache, fatigue, decreased energy, decreased alertness, drowsiness, depressed mood, difficulty concentrating, and irritability. The incidence of headache was 50% and the incidence of clinically significant distress was 13%. The onset of symptoms occurred 12–24 h after abstinence, with peak intensity at 20–51 h and for a duration of 2–9 days. Abstinence from even low doses (e.g. 100 mg day^{-1}) produced symptoms. Unfortunately, this review was selective and studies which suggested a different view of caffeine withdrawal were excluded. In addition, there was no attempt to distinguish between negative effects of withdrawal and positive effects of caffeine. For example, certain studies are interpreted in terms of negative effects of caffeine deprivation when they could actual be interpreted in terms of positive effects of caffeine ingestion. This review does not look at the details of the studies and, as stated above, many effects of caffeine deprivation are no longer significant when adjusted for multiple statistical tests.

Recent research has demonstrated that acute caffeine abstinence produces changes in cerebral blood flow

velocity, EEG, and subjective effects. These vascular effects of caffeine withdrawal are clearly very different from behavioral effects of caffeine which are thought to reflect changes in a variety of neurotransmitter systems. Astrid Nehlig and colleagues present evidence that in animals caffeine does not trigger metabolic increases or dopamine release in brain areas involved in reinforcement or reward. A single photon emission computed tomography (SPECT) assessment of brain activation in humans showed that caffeine activates regions involved in the control of vigilance, anxiety, and cardiovascular regulation but did not affect areas involved in reinforcement and reward.

TOLERANCE

Developing tolerance is a hallmark of substance abuse and dependence. In adults, caffeine-induced tolerance has been shown for some, but not all, outcomes and only in a sub-set of consumers.

CROSS-SENSITIZATION

Cross-sensitization is the process by which taking one drug enhances the response to other drugs with the same neurobiological mechanisms. It has been suggested that caffeine may increase sensitization to nicotine although the correlation between caffeine use and smoking may reflect the faster metabolism of caffeine by smokers. There is no clear relationship between caffeine use and cocaine in humans and if anything, cocaine users are less likely to consume caffeine than non-cocaine users.

CAFFEINE AND CHILDREN AND ADOLESCENTS

Ingestion of caffeine from naturally occurring sources has been largely restricted to adults but it is now added, sometimes in large quantities, to drinks that are consumed by children. Our knowledge of the effects of caffeine on the behavior of children needs to be extended by further research. The current position on this topic can be briefly summarized as follows.

Older studies of the behavioral effects of caffeine on children have shown similar effects to those observed in adults. Effects in children are often smaller than those observed in adults which may reflect the smaller doses consumed.

It is generally agreed that caffeine intake by pregnant women should be kept at a low level (below 200 mg) because of the possible impact on birth problems and reduced body weight of the child. However, there is no evidence showing that caffeine consumption during pregnancy or childhood influences brain development.

A recent review has shown that caffeine containing drinks are now regularly consumed by children. Indeed, some caffeinated products are even marketed to children as young as 4 years old. Our knowledge of effects of caffeine on children is very limited and further research is needed in the area because children may be more sensitive to negative effects of caffeine than adults. This research should examine possible caffeine dependence and also caffeine intoxication. Caffeine intoxication is characterized by the following symptoms: restlessness, nervousness, excitement, insomnia, flushed face, diuresis, and gastrointestinal complaints. It is likely that low consumers of caffeine, such as children, may experience caffeine intoxication following consumption of a high dose (as found in some energy drinks). No empirical studies have been conducted to examine whether children and adolescents develop tolerance to the effects of caffeine. Surveys suggest that 41.7% of teenagers reported tolerance to caffeine and 77.8% reported symptoms of withdrawal. Consumption of caffeinated soft drinks is also associated with poor diet, excess weight, and dental caries.

There is a growing literature that suggests that caffeine use in adolescents and young adults is associated with impulsivity, risk taking, and sensation seeking. Unfortunately, due to the correlational nature of these studies, it is not possible to determine the direction of causality.

ENERGY DRINKS

Energy drinks represent the fastest growing sector in the beverage industry. These drinks often contain five times the amount of caffeine as soft drinks and may also contain taurine, riboflavin, pyridoxine, and various herbal derivatives. Most energy drinks also contain sugar in an amount that exceeds recommended daily allowances. Studies of the effects of energy drinks on behavior confirm that they increase alertness and attention, improve simulated driving when sleepy and can reduce sleepiness in night workers. However, energy drinks that also contain alcohol (6% by volume) have been shown to impair a global measure of cognitive functioning. Energy drinks have also been shown in laboratory studies to increase heart rate and blood pressure.

Energy drink consumption can lead to caffeine intoxication especially in children. Deaths attributed to energy drink consumption have been reported in Australia, Ireland, and Sweden. Health care providers

report the following effects after consumption of energy drinks: dehydration, accelerated heart rates, anxiety, seizures, acute mania, and strokes. The risk of caffeine intoxication may be greater for energy drinks than for other sources of caffeine due to inadequate labeling, advertising, and the consumer demographics.

Energy drinks are often combined with alcohol to increase the positive effects of alcohol ingestion and counteract the depressive effects. This can lead to increased alcohol intake and an increase in adverse events due to alcohol. Indeed, combining energy drinks with alcohol gives the person a false sense of control. Recent research has investigated the extent to which energy drink consumption was a risk factor for alcoholism. The results of a study of over 1000 university students showed that weekly or daily energy drink consumption was associated with more frequent and greater consumption of alcohol. This, of course, could be due to alcohol consumption influencing energy drink consumption rather than the other way around.

CONCLUSIONS

In conclusion, there are many beneficial effects of caffeine and negative effects are restricted to consumption of high doses by susceptible individuals. Caffeine is almost certainly the most widely used drug of dependence in the world. Despite this, the evidence of morbidity associated with caffeine consumption is slight. Research on caffeine tells us little about the harmful effects of drugs of dependence and shows that caffeine dependence *per se* is not a problem.

SEE ALSO

Tobacco, Food Addictions, Khat Addiction, The Biopsychosocial Model of Addiction, Tolerance and Withdrawal

List of Abbreviations

OTC	over-the-counter
SPECT	single photon emission computed tomography

Further Reading

Dews, P.B., O'Brien, C.P., Bergman, J., 2002. Caffeine: behavioural effects of withdrawal and related issues. Food and Chemical Toxicology 40, 1257–1261.

Fredholm, B.B., Battig, K., Holmen, J., Nehlig, A., Zvartau, E.E., 1999. Actions of caffeine in the brain with special reference to factors that contribute to its widespread use. Pharmacological Reviews 91, 83–133.

Glade, M.J., 2010. Caffeine – not just a stimulant. Nutrition 26, 932–938.

Goldstein, E.R., Ziegenfuss, T., Kalman, D., et al., 2010. International society of sport nutrition position stand: caffeine and performance. Journal of the International Society of Sports Nutrition. 7 (1), 5.

James, J.E., Keane, M.A., 2007. Caffeine, sleep and wakefulness: implications of new understand about withdrawal reversal. Human Psychopharmacology 22, 549–558.

James, J.E., Rogers, P.J., 2005. Effects of caffeine on performance and mood: withdrawal reversal is the most plausible explanation. Psychopharmacology 182, 1–8.

Juliano, L.M., Griffiths, R.R., 2004. A critical review of caffeine withdrawal: empirical validation of symptoms and signs, incidence, severity and associated features. Psychopharmacology 126, 1–29.

Ker, K., Edwards, P.J., Felix, L.M., Blackhall, K., Roberts, I., 2010. Caffeine for the Prevention of Injuries and Errors in Shift Workers (Review). The Cochrane Collaboration. Wiley, Chichester.

Lara, D.R., 2010. Caffeine, mental health, and psychiatric disorders. Journal of Alzheimer's Disease 20 (Suppl. 1), S239–S248.

Lieberman, H.R., Tharion, W.J., Shukitt-Hale, B., Speckman, K.L., Tulley, R., 2002. Effects of caffeine, sleep loss, and stress on cognitive performance and mood during U.S. Navy SEAL training. Psychopharmacology 164, 250–261.

Reissig, C.J., Strain, E.C., Griffiths, R.R., 2009. Caffeinated energy drinks – a growing problem. Drug and Alcohol Dependence 99, 1–10.

Smith, A.P., 2005a. Caffeine at work. Human Psychopharmacology 20, 441–445.

Smith, A.P., 2005b. Caffeine. In: Lieberman, H., Kanarek, R., Prasad, C. (Eds.), Nutritional Neuroscience. Taylor & Francis, pp. 335–359.

Smith, A.P., 2009. Caffeine, cognitive failures and health in a non-working community sample. Human Psychopharmacology: Clinical and Experimental 24, 29–34. http://dx.doi.org/10.1002/hup.991.

Temple, J.L., 2009. Caffeine use in children: what we know, what we have left to learn, and why we should worry. Neuroscience and Biobehavioral Reviews 33, 793–806.

CHAPTER

79

Food Addictions

Mark S. Gold, Richard L. Shriner

College of Medicine, University of Florida, Gainesville, FL, USA

OUTLINE

DEFINITION

Establishing a workable definition for food addiction is of paramount importance to understanding one of the world's most pressing epidemics, obesity. This definition must also lend greater definition to other forms of food addiction such as binge eating, bulimia, anorexia, and other maladaptive relationships with food. If food is to be seen as a putative agent for addiction and obesity, then we must simultaneously entertain it in a framework not unlike other substances of addiction. A working definition of food addiction should identify a distinguishing set of characteristics that sets it apart from normal eating. It must clearly outline a distinct pattern of behavior that is truly unhealthy (i.e. maladaptive) over space and time. Such a definition is offered below. Importantly, food addiction (like other addictions) involves a dual biasing valence of both perception (i.e. a highly valued [rewarding] substance) and behavior (i.e. craving/dependency), which reliably sets up the abuser and food addict to predictably suffer a set of unhealthy outcomes.

In terms of *Diagnostic and statistical manual of mental disorders 4th edition text revision* (DSM-IV-TR), researchers have argued that food (like gambling) is indeed an addictive substance. They argue that food fulfills the seven DSM-IV-TR criteria of substance dependence. This includes three or more of the following seven criteria realized over a period of at least 12 months: (1) tolerance, (2) withdrawal, (3) the substance is taken in larger amounts for a longer duration than initially intended, (4) attempts are made to cut back on the substance, (5) excessive time is spent pursuing, using, or recovering from the substance, (6) there is curtailment of other activities because of use of the substance, and (7) there is the continued use of the substance despite adverse consequences that are caused by the substance.

In terms of practical clinical interventions, we offer the following working definition for food addiction:

> *Food addiction represents a pervasive and enduring pattern of both food perception and food-related behavior (leading to either excessive food ingestion or aversion) whose dual valence (i.e. perception and behavior) biases interaction with food in harmful and unhealthy ways. Such a biasing and unhealthy valence toward food continues, despite knowledge of its harmful consequences. Food addicts usually present both a tolerance (i.e. a need to increase participation in their harmful relationships with food over space and time) as well as a form of withdrawal (i.e. an inability to escape their addiction with food without suffering undue anxiety, craving, or other adverse neurochemical reactivity [which may include depression or anger]) when deprived of access to addictive foods. This latter emotional and behavioral reactivity must reliably occur during efforts to either alter or disrupt the food addict's harmful and maladaptive pattern of eating.*

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00079-6

The advantage of this definition is that unlike other forms of addiction, ultimate food sobriety (defined later) is clinically understood to be neither absolute nor static, but rather relative and dynamic. Using this sort of dynamic conceptualization of living in balance with food (instead of being captured by its pathological valences causing fixed weight gain, pathological weight flux or loss) the term Living with Food seems justifiable and clinically useful. This living or healthy coexistence with food involves a state of dynamic interaction that implies neither dependency nor addiction, but rather food sobriety and health. Such sobriety connotes a new, dynamic, and more resilient relationship with food that allows for a healthy acquisition and maintenance of weight over space and time. This is done by balancing those valences that would otherwise cause a person to tend toward pathological patterns of eating. Such pathological patterns having genetic, anatomic, neurochemical, psychological, and environmental roots and triggers, are described under the rubric of disordered eating. These disorders include a spectrum reaching from obesity to binge eating to bulimia and then on to anorexia. This newly formulated, dynamic definition of food addiction paves the way for the study of the entire spectrum of disordered eating under a common defining schema, from excessive weight gain (obesity), to periodic weight gain (binge eating and/or bulimia), to excessive weight loss (anorexia).

MAGNITUDE OF THE PROBLEM

If we assume that food addiction plays a potent role in obesity (an assumption that now appears quite plausible) and obesity is the single most common cause of type 2 diabetes mellitus (T2DM), then food addiction may be the most singularly important addiction of our time. In this way, food addictions will continue to vex scientists, minions of public health policy, insurance companies, the Federal Government, and tax payers who will be asked to cover an ever-escalating health care budget as more graduate into the ranks of obesity. Here are some sobering statistics.

By 2018, more than 43% of US citizens will be obese. Morbid obesity is one of the fastest growing disease states. The National Council on Compensation Insurance found that

- obese claims were roughly three times more expensive;
- added treatments related to obesity can balloon cost differences by as much as 30 times;
- workers who are morbidly obese filed 45% more claims, missed eight times more work, and experienced medical costs five times higher than nonobese counterparts.

The world picture in terms of obesity and food addictions (especially for high carb/high fat food (cafeteria style or highly palatable food)) is no better. Rates of obesity have tripled in the last 20 years in the developing world with 10% of the world's children currently overweight or obese. Countries facing the greatest challenges include the Middle East, Pacific Islands, Southeast Asia, and China. Amongst major world powers, McDonalds (in 6 out of 11 countries) is the favorite fast food chain, with China favoring Kentucky Fried Chicken (KFC). In terms of obesity, the body mass index (BMI) for US children at the 95th percentile is below that in China. Australian and UK women are rapidly approaching the BMIs found in US women. Amongst children in the major countries of the United States, United Kingdom, Australia, and China, Chinese children have gained the marginally highest, some 5-point increase in BMI since 1991. Three things are now self-evident: (1) the major type of diabetes is T2DM, (2) T2DM is largely preventable through weight loss, and (3) T2DM is the number one cause of death in the United States. Hence, the cycle of obesity, diabetic misery, and death tells us food addictions play an ever more critical role in the survival of our species. This cycle goes on and on unabated. The *Morbidity and Mortality Weekly Report* (MMWR) August 2010 report found that rates of obesity have increased by 1.1% since 2007, negating any prior reports of a leveling off of the obesity epidemic.

SUPPORTIVE EVIDENCE

Supportive evidence for food addiction comes from the following lines of research:

- Genetic
- Neuroanatomic/neuroimaging
- Neurochemical
- Stimulus/reward paradigms
- Psychological research
- Clinical case examples

Genetic

Genetic contributions to food addiction seem plausible given the fact that addiction has an estimated heritability of 50–70%. Various researchers have described how such genetic factors are at play in the dysmorphic syndromes of Prader-Willi, Cohen, Carpenter, and others, as well as leptin/receptor mutation, β3 AR mutation and the overexpression of neuropeptide Y (NPY). Each one of these genetic syndromes

results in obesity and can be seen to exacerbate food addictions in the individuals that are so afflicted. As described later in the neurochemical section, dopamine plays a critical role in addiction. The neurogenetics of dopaminergic receptor supersensitivity and brain reward circuitry continues to emerge. For example, a plethora of animal models are being used to elucidate the role that genetics plays in both animal and human vulnerability to addiction in terms of their responsiveness to food or cocaine.

Neuroanatomic/Neuroimaging

Neuroanatomic and neuroimaging studies that are germane to understanding food addiction continue to be developed. Investigators have shown how (using positron emission tomography [PET] and [^{18}F]fluorodeoxyglucose [FDG]), morbidly obese patients have increased metabolism in the somatosensory cortex. They have also showed using PET and FDG in normal weight individuals how increased dopamine increases metabolism in the orbitofrontal cortex and is directly involved with the perception of hunger and desire for food. Using group averaged images of [^{11}C] raclopride PET scans, these same researchers have been able to show that obese rats have marked diminution of dopamine transporter levels compared with nonobese controls. This suggests downregulation of the dopamine (D2/D3) receptor function in obese subjects. This has been replicated in humans. Over the last decade, we now have a greater appreciation for the science of food addiction, including two distinctive food addictive pathways involving: (1) the homeostatic and (2) the reward (or hedonistic) circuits of the brain. Functional magnetic resonance imaging (fMRI) studies have revealed how food craving is activated specifically through three distinct areas of the central nervous system (CNS): hippocampus, insula, and caudate.

Neurochemical

Shadowing the progress of understanding food addictions through neuroimaging of the brain, our neurochemical understanding of food addition and disordered eating continues to be elucidated. In terms of the recent worldwide obesity epidemic, this includes how appetite control and energy balance have been adversely affected by large amounts of high energy-dense foods (e.g., high fructose corn syrup [HFCS]), increased exposure to powerful food cues (e.g. television advertising), together with sedentary lifestyles to create an imbalance between repletion signals versus reward-driven brain inputs leading to obesity. Researchers are focusing on the processes of bingeing, withdrawal, craving, and sensitization that can be used to define food addiction (as well as addictions in general). Again, investigators have found the three key neurotransmitters (dopamine, opioids, and acetylcholine) play a pivotal role in the development and maintenance of food addiction through time. Currently, a major focus of study involves the role that intermittent excessive carbohydrate/cafeteria style (fast food) ingestion may have on binge eating. Specifically, rats exposed to excessive amounts of sugar, which were then denied access to such sugar, tended to engage in binge eating over time, versus rats that were given continuous access to such high carbohydrate food (see below).

Continuing along this line of investigation, binge eating is also associated with stress. Together with an intermittent exposure to palatable foods (e.g. eating fast foods on an intermediate basis), the evolution of binge eating resulting in the exacerbation of weight seems to reliably involve opioid chemistry interaction within the nucleus accumbens of the brain, along with specific dopamine inputs. Increasingly, animal models for the neurochemistry of food addiction (especially involving sugar) are proving the existence of a craving–bingeing–withdrawal–craving cycle. Researchers in the area of food addiction have described how sugar bingeing works through the combined inputs of dopamine interacting with acetylcholine to affect the net output of γ-aminobutyric acid (GABA), all within the nucleus accumbens. Then, this GABA neurochemical traffic goes on to mediate either approach (increased appetite) or avoidance (decreased appetite). In this way, acetylcholine may work to oppose the stimulating effects that dopamine has on appetite. Interestingly, these same researchers found that although both sugar and fat can be associated with bingeing, only sugar was associated with the classic signs of withdrawal, as one would expect from a truly addictive substance.

Stimulus/Reward Paradigms

Stimulus/reward paradigms involving neurobiological pathways and putative neurochemical agents for food addictions are now solidly in play. Again, investigators have shown the significance of addiction to food and brain reward systems. They argue that opioid-based neuropharmacology may underlie the mechanisms behind our current over consumption of palatable (i.e. high carb/high carb, high fat) foods. In general, most research points to the fact that opioid stimulation increases the pleasurable aspects of food, while opioid antagonists do the opposite. At least for carbohydrate craving, opioids may decrease oxytocin-mediated satiety (oxytocin normally turns off carbohydrate

cravings). Thus, opioids may help to mediate the food reward value of carbohydrates. This includes sugar-induced upregulation of μ-opioid and D1 dopamine receptors, along with an increase in release of dopamine in the nucleus accumbens. In fact, in rat models, sucrose is often chosen over cocaine. As we just alluded to above, an emerging body of research suggests that sugar can be highly addictive, causing both craving and bingeing in animal subjects who are denied access to it. To be sure, studies of rats who have been exposed to intermittent schedules of high-dose sugar followed by its sudden removal, can display frank withdrawal (i.e. increased motility, nervousness, etc.) after the opioid antagonist naloxone is administered to them. In humans, opiate-dependent individuals on methadone maintenance report higher consumption of sweets than control subjects.

Lastly, in terms of stimulus/reward paradigms, food addictions (including binge eating and obesity) highlight that time and space may play an important part in obesity and food addictions, as we suggested in our definition of food addiction given at the end of the last section. Specifically, the effects of some food stimulus early in our lives may have an effect later in life vis-à-vis the way we live with food. A recent study was able to demonstrate that if mice were exposed early in their development to high carbohydrate chow, although they did not immediately gain weight, later on in their development they did show increased obesity (over controls) when these same mice were given access to highly palatable food. Thus, early exposure to sugar may have an enduring effect (and/or influence upon) our capacity to suffer from obesity later in life.

Psychological Research

Another area of evidence that supports food addiction involves psychological research. This involves whether psychological mechanisms play a part in food addictions, albeit as modified by genetic and epigenetic mechanisms. Research is now giving us an interesting glimpse into the possible inputs that psychobiological factors such as reward sensitivity and impulsivity may have on food addictions and obesity/eating disorders. Investigators have recently outlined how genetic inputs to dopamine function (hence food motivation) are translated by way of polymorphisms of the D2 receptor (Taq1A), which may have direct impact on obesity and food addiction. The Taq1A+ allele has reduced brain dopamine function compared with the A1 allele since those with the Taq1A+ allele have a 30–40% reduction in D2 receptor density in the striatal region. This has been related to reduced mesolimbic brain dopamine, therefore, a deficient ability to experience natural reward. Given that patients with Taq1A+ allele function have a hypoactive reward system, these same individuals may be susceptible to low food stimulus/reward functionality. Thus, these same individuals may need to consume larger amounts of food to achieve the same measure of reward that those with A1 allele function require. This leaves Taq1A+ allele cohorts at risk for over ingestion and obesity. At least for some individuals with a high BMI, this has been confirmed. The μ-opioid receptor gene (OPRM1) has been extensively studied for drug abuse. This distinct genotype may be higher in binge eaters, but perhaps not in obese individuals without binge eating disorder.

Other studies have examined the psychological effects on obesity and food addiction that carbohydrates may have on female subjects suffering depression. As we mentioned in our definition at the end of the last section, individuals suffering from food addiction may undergo both anxiety and/or depression when they face the prospect of food withdrawal. This, of course, forms the withdrawal component within the definition of food addiction. Investigators have been able to show that after inducing dysphoric mood in overweight female subjects, carbohydrate beverages (but not taste-matched protein substitutes) were able to effectuate an antidepressant effect. In other words, carbohydrates (in contrast to matched taste proteins) were able to assuage the carbohydrate cravings these dysphoric females were suffering from, suggesting that carbohydrate ingestion medicates mildly dysphoric mood.

Clinical Case Examples

The last area of evidence that supports the existence of food addiction involves clinical case examples in general medicine; specifically, how stress and stress-related medical disorders may be involved with carbohydrate cravings and addictions. As we stated earlier, stress plays a crucial role in animal models of binge eating. To this end, if rats are exposed to food restriction alone, they did not tend to overeat. But if one adds the stress of foot shock trauma, these same rats tend to significantly over ingest, especially palatable foods high in carbohydrates and fat. In studies showing this effect, rats needed a primer of at least three restriction/refeeding cycles (see earlier discussion on intermittent exposure to sugar and bingeing), as would be the case in humans for those trying to diet by intermittently restricting their food intake. This is, of course, the same phenomenon we discussed earlier in the rats who were intermittently given access to high carbohydrate chow, followed by restriction. Again, such animals were shown to be at high risk for the eventual development

of binge eating. High stress reactivity binge type eating has also been shown in humans.

Food and drugs have cross tolerance in the laboratory and food, especially sugar, is used to calm withdrawing cigarette smokers and addicts. Actual clinical case examples of how food addiction and stress may play a part in everyday medicine comes from research involving models of stress, allostasis, and chronic diseases. The literature on allostatic load refers to the consequences of sustained or repeated activation of mediators of allostasis. Allostasis (literally other balance) is invoked when specific effectors are activated because organismal homeostasis (literally similar balance) is threatened. An example of such an effector might be an interleukin called IL-6, which is a proinflammatory marker in the body. Such inflammatory markers are released when the body undergoes stress. Such stress-induced inflammatory markers tend to be associated with a simultaneous immune response that forces the egress of cortisol from the adrenal gland (perhaps to counteract such effectors). In turn, cortisol can go on to stimulate appetitive function since it has an orexigenic influence on the hypothalamus. Living within this environment, the organism is predisposed to increased hunger, dopamine release, and (as we have learned earlier) increased motivation to seek out food. Thus, stress can through allostatic load, lead to both direct and indirect reactivities in the body that then go on to induce disturbances of appetite and invariably weight gain. In fact, waist to hip ratio (a measure of obesity) as well as glycosylated hemoglobin (a measure of excess blood glucose) are two important indicators of allostatic load.

A real-world clinical example of how food addiction, allostatic load, allostasis, and orexigenia might find its way into our lives is found in the disorder called metabolic syndrome. This disorder (again really a syndrome) exists when a patient suffers from at least three of the following: (1) hypertension, (2) hypertriglyceridemia, (3) decreased high-density lipoprotein fraction, (4) hyperglycemia; and (5) obesity. In the case of food addiction, continued cravings for especially high carbohydrate substances continue to inflame body effectors (e.g. IL-6, tumor necrosis factor (TNF), etc.) ultimately causing the exhaustion of insulin-producing cells in the pancreas leading to the number one cause of diabetes, T2DM. Regardless, these sorts of metabolic imbalances can lead to obesity and diabetes, which increase the vulnerability of an individual to stress, which creates a vicious cyclical mechanism between stress, food addictions, obesity, diabetes, leading to more stress. As if this were not enough, stressful life events are known to accelerate the course of diabetes in adverse ways. Since diabetes costs approximately 2.5 times more than any other disorder to treat, diabetes, obesity, food addiction, and runaway health care costs remain inextricable from one another.

T2DM AND WHY FOOD ADDICTIONS MATTER

A Model of Pro-inflammation and Energy Misalignments

The reason food addictions matter, at least in terms of their impact on obesity and medical economics in general, is that obesity is so insidiously and directly associated with T2DM, the number one killer in the United States, if not soon the entire world. Certain endocrinologists have pointed out that food (in particular, HFCS) and other substances of addiction (e.g. ethanol) have many metabolic perturbating and allostatic (see earlier discussion) disease-contributing effects in common. These same endocrine researchers convincingly argue that both HFCS and ethanol (EtOH) can contribute to hyperlipidemia, hepatic inflammation, and hepatic (if not also skeletal) insulin resistance. This leads to increasing feedback-mediated hyperinsulinemia (i.e. to compensate for insulin receptor resistance (insensitivity) in order to preserve organismic metabolic integrity) which eventually exhausts β-cell pancreatic function, often leading to T2DM, if not eventually type 1 diabetes in selected individuals.

When it comes to food addiction and its impact on obesity, research is teasing out a common thread between foods of abuse and their impact on several key metabolic, endocrine, and neuroendocrine pathways. These processes and pathways seem to involve the following unraveling of healthy hepatic metabolic eustasis (meaning literally good balance):

- exotic (not previously available to *Homo sapiens*) nutrient energy excess (e.g. HFCS);
- ensuing perturbations of hepatic metabolic function;
- stimulation of proinflammatory species leading to hepatic insulin resistance;
- overwhelmed metabolic pathways that stimulate hepatic gluconeogenesis;
- overwhelmed metabolic pathways that stimulate hepatic dyslipidemia;
- resultant hypersecretion of insulin which favors development of T2DM;
- resultant hypersecretion of insulin which favors fatty acid storage;
- skeletal muscle insulin resistance/hyperlipidemia that decreases energy expenditure;
- net weight gain due to decreased fatty acid mobilization from adipocyte to muscle;
- long-term leptin resistance caused by hyperlipidemia causing further hyperphagia;

- hyperuricemia from excess ADP breakdown causing NO causing blood pressure;
- all of the above contributing to metabolic syndrome and preventable death.

The most potent exotic source of nutrition that has arrived on the landscape of the modern *Homo sapiens* is HFCS. This inflaming source of energy causes several specific perturbations in hepatic function that have been identified by key researchers in the study of endocrinology and obesity. These same researchers have done a masterful job of pointing out the similarities between the addictive substrate EtOH and HFCS. Specifically, how each of these two addictive substances can cause proinflammatory changes, possible reactive oxygen species (ROS), dyslipidemia, and ensuing insulin resistance. EtOH, by altering GABA and opioid transmission within the ventral tegmental area and central areas of the amygdala, activates dopamine neurotransmission. As we mentioned earlier, the stimulation of both opioid and dopamine receptor function is the same thing that carbohydrates (including fructose) are capable of doing and is a prerequisite for defining a substance as addictive.

In this way, sugar, EtOH, HFCS, and other potential addictive substances affect not only peripheral organs, but central nervous nuclei as well. Collectively, these and other addictive substances have been shown to cause the stimulation of a host of proinflammatory substances. Such substances are known to be perturbated in an allostatic fashion, which collectively reflect a state of metabolic dysresiliency called the metabolic syndrome. Woven into this syndrome is diabetes, obesity, hypertension, and dyslipidemia, all reflecting a state of hyperoxidative distress. During such states of oxidative organismic dysfunction, adiponectin levels drop while IL-6, TNF, and other proinflammatory markers are increased. In terms of food addictions, it is this state of metabolic dysfunction or stress (what we have chosen to call metabolic dysresiliency) that identifies a state of abuse for the obese food addict.

The qualifier (and/or proinflammatory quantifier) that outlines a state of abuse in terms of food addiction is found in the net energy expenditure variables that lead to either metabolic resiliency (when energy intake and energy expenditure are resiliently matched to maintain an adaptive dynamic weight called eubaria) or when they are in nonresilient allostasis reflecting a maladaptive weight that is either hypobariac (e.g. anorexia) or hyperbariac (e.g. obesity). A heterobariac weight would be found in the bulimic individual who although maintaining a normal weight is nonetheless still living in a state of dysresiliency that betrays underlying metabolic distress. This is why these individuals often suffer from neuropsychiatric disease states that reflect hypothalamic–pituitary–adrenal axis abnormalities (e.g. increased level of corticotropin-releasing factor) such as depression and/or anxiety.

At least for the hyperbariac conditions such as simple overweight and obesity, food addictions that lead to a state of abuse invariably lead to metabolic dysresiliency and dysfunction. These metabolic dysresilient states are often, unfortunately, self-sustaining since they tend to disturb and upset the normal balance of anorexigenic/orexigenic endocrine feedback (e.g. food addictions causing obesity, block leptin, hence the body's normal anorexigenic responsiveness, etc.).

With reference to the term sobriety, it follows that food sobriety would reflect a state of eubaria and/or metabolic resiliency. To that end, we can define the term food sobriety in the following way:

> *Food sobriety is a dynamic state of interactivity with food that maintains a healthy balance of resilient metabolic function that is free of significant symbolic (relational) or hedonistic (addictive) skew.*

In terms of obesity and T2DM, such a healthy balance of metabolic function or metabolic resilience can be defined as a ratio of insulin to glucagon that is kept in a dynamic (pulsatile) adaptive flux. This flux is necessary for normal hepatic and pancreatic function to work in healthy synergy with one another. This is reflected by a pattern of food ingestion that creates a normal adiponectin and a low proinflammatory oxidative state. To be sure, research has found that the normal healthy pancreas releases insulin in a dynamic, somewhat chaotic fashion that is pulsatile and continuously adaptive, in a nonlinear fashion. This is, of cours, a state much different than that emulated by taking pills or injections to regulate insulin via pharmaceuticals. Even insulin pumps cannot possibly totally emulate this nonlinear continuously adaptive, unpredictable moment to moment resilient flux of insulin versus glucagon that the native healthy pancreas is able to engineer. Therefore, all present and future medical interventions must emphasize the importance of salvaging and restoring native hepatic and pancreatic functioning whenever possible. This, of course, is not always possible, often making permanent and ongoing pharmacotherapy mandatory, if not lifesaving.

Clinical Presentation

The clinical presentation of food addiction is obviously both global and omnipresent. Food addictions start with the way we offer nutrition and take care of our children and end in nursing homes and hospitals, which are, even now, overwhelmed by growing numbers of patients dying of metabolic syndrome and T2DM. We need to do a better job of early intervention,

translational medicine, and more aggressive treatment in terms of patient and physician information sharing, dietary intervention, exercise incentivization, stress/wellness intervention, and psychobiological/psychological evaluation and treatment.

The most important initial assessment in the diagnosis of patients suffering from food addiction is the history and physical examination. At our obesity clinics at the University of Florida, patients are asked to fill out an extensive bariatric data packet. This packet includes background data on current stresses, past stresses (including childhood trauma), past medical and surgical history, past psychiatric history, participation in either Overeaters Anonymous, Alcoholics Anonymous, Weight Watchers, or other fellowship programs, past and current medication, a figurine analysis of the patient's perceptions of their body image, a patient-generated narrative of their self-image, a thorough life–weight profile beginning in childhood up until the present, a canvassing of types of foods eaten, frequency and amount, assessment of binge eating and other eating disorders (e.g. laxative, diuretic abuse, self-induced vomiting, etc.), a list of all diets and diet supplement/diet pill pharm trials, a sleep apnea assessment, inquiry about reasons patient feels they are overweight, and a series of brief psychiatric rating scales. The packet is quite easy to fill out and provides our staff with invaluable data that might otherwise not be assayed.

Next, patient's blood and urine are carefully assessed, including a complete fasting metabolic panel, complete blood count, thyroid functions, fasting insulin, and C-reactive protein. Each patient's QUICKI score is calculated to assess a raw estimate of insulin resistance. In addition, we may or may not obtain IL-6 and adiponectin titers . After obtaining suitable blood and urine chemistries, weight, electroencephalogram, etc., to assay for disturbances in metabolic function, all our patients are given a cognitive assessment, which helps us discern exactly how they negotiate health care information under stress.

Using a cognitive assessment tool allows today's modern bariatrician to interact with patients in terms of their own biases and unique modes of interacting with food. These biases can often be shown to interfere with a patient's capacity to overcome their food addiction. Patients should be taught how certain foods, along with their time/space ingestion patterns (see earlier discussion on intermittent schedules of feeding) can lead to hedonistic problems with food, which can then lead to binge eating and other maladaptive ingestion patterns. Finally, patients should be taught how ingesting large quantities of carbohydrates (especially HFCS) can create metabolic problems with food sobriety given their capacity to upset insulin to glucagon ratios and damage hepatic metabolic function.

Giving patients the capability of focusing upon their relationship problems with food sobriety, cognitive tools (such as the IGS-rooms tool we have developed at the University of Florida) helps patients identify their maladaptive relationship patterns not only with food, but also with their relationships with others (at home, at work, and at play). As we mentioned in the section involving the pivotal role that stress plays in maladaptive eating, without such a cognitive assessment and enrollment in group fellowship intervention, patients tend to remain in maladaptive relationships with food. This may be why the vast majority of diets fail over time. These patients, without adequate medical translation tools, have not engineered a complete transformation in their metabolic, addictive, and relationship interactions with food. This makes food sobriety difficult, if not impossible.

After completing a cognitive assessment, patients should undergo a thorough dietary evaluation that helps to quantify and qualify the exact ratios of fats, proteins, and carbohydrates they are ingesting in a typical day. Effort should be made to assist the patient in minimizing their exposure to HFCS as well as large quantities of EtOH for reasons clearly spelled out above. This is exactly why food addiction programs fit comfortably alongside alcohol and other polysubstance abuse interventions. Next, any underlying neuropsychiatric dysfunctions should be addressed via pharmacological and/or psychotherapeutic intervention.

Whenever possible, patients should be offered group fellowship interventions. Such group interventions can explore maladaptive patterns of thinking and relating to food. These patterns are then reset into more resilient or adaptive complexes. The group process can then reinforce through group fellowship an enduring participation and validation of more resilient and adaptive interactions with food. Since much of the literature on food addictions seems to support the notion that high carbohydrate/cafeteria style foods (especially when periodically curtailed in repetitive attempts at dieting) create more binge eating than weight loss, programs should generally support continuous feeding throughout the day. The overall goal with patients should be to achieve a relatively low carbohydrate/reasonably high protein/reasonable medium chain fatty acid diet. Supplementing a patient's diet with low carbohydrate/reasonably high protein whey protein shakes also seems worthwhile. Rapidly absorbed proteins such as whey protein have been shown to help mitigate hunger. Such a low carbohydrate/high protein caloric source does not tend to aggravate insulin production. Finally, in order to mitigate ongoing inflammation, omega 3 fish oil augmentation would seem advisable in patients wishing to mitigate the proinflammatory responses secondary to their obesity.

The typical food addict is often a patient who has T2DM and/or a patient who is manifestly obese. This includes bariatric surgical patients (who usually have BMI >40 kg m^{-2}). However, clinicians will also discover food addictions in heterobariac patients (defined above), whose BMI is between 17.5 and 25 kg m^{-2}. Obviously, such patients may suffer from either borderline anorexia nervosa (i.e. just above 17.5 kg m^{-2}) or bulimia (with a normal BMI). In addition, binge-eating patients are obviously found nested within general obese patient populations, representing (conservatively) some 5% of these patients. It has been suggested by various investigators that binge eaters have more intrinsic neuropsychiatric disturbances than those who suffer from non–binge-eating obesity. Others who may suffer from food addictions include patients with polycystic ovaries syndrome, chronic pain (especially those who suffer from concomitant problems with opioid addictions), chronic pulmonary disease, various plegic or paretic disorders leaving them wheelchair bound, patients who are cardiovascular compromised, female rather than male (worldwide, females tend to be more obese than their male counterparts), postpartum patients, etc. Interestingly, certain investigators have shown that morbid obesity is often not associated with alcoholism. Overeating may compete with alcohol for brain reward sites, making alcohol ingestion less reinforcing.

Current Treatments

The treatment of food addiction is complicated by the diagnosis. Are there one or many types of food addictions? Is fat preference and overeating a different or related form of sugar and desert-loving food addiction? Only additional research will tell. Multimodal treatment is the norm, extending from group fellowship intervention, to individual psychotherapy, to strategic diet targeting, to focused exercise to bariatric surgical intervention to specific target psychopharmacology. The various antiobesity pharmacologic reagents (A-OPRs) are as diverse as the organs these drugs seek to influence to carry out their anorexigenic response.

In terms of the CNS, A-OPRs include bupropion, the serotonin noradrenaline reuptake inhibitor (SNRI) antidepressant (Wellbutrin). This antidepressant effectuates weight loss via its pro-opiomelanocortin (POMC) activator properties. More recently, naltrexone (an opioid receptor antagonist) has been coupled with bupropion to augment bupropion's POMC activator potency. The anticonvulsants topiramate (Topomax) and zonisamide (Zonegran) have shown some success. Phentermine (Adipex) is a stimulant but only has US Food and Drug Administration (FDA) indications for short-term use only. Unfortunately, patients usually quickly regain weight after discontinuation of the drug. Sibutramine (Meridia) is a central norepinephrine/serotonin reuptake inhibitor whose use is limited by possible hypertensive side effects. Together with orlistat (see below), sibutramine is the only FDA-approved medication for weight loss. Rimonabant (Acomplia) is an endocannabinoid type 1 inhibitor that was taken off the market because of significant neuropsychiatric side effects including suicidal ideation. More recently, DOPA 3 agonists have been in development and may have potential value in impulsive behavioral diatheses such as cigarette addiction. Some investigators are now considering these same reagents in the treatment of binge eating.

In terms of the gut, orlistat (Xenical) is a lipase inhibitor that impairs the digestion of fat causing some degree of malabsorption and weight loss. Its use is limited by the emergence of fatty stools and bowel leakage. Experimental A-OPRs in various phases of development include obinepitide (a dual analogue of PYY3-36 from the distal gut's L cells and PP from the F cells of the pancreas), GLP-1, OXM (Oxyntomodulin, which contains glucagon and is a potent central acting anorexigenic agent), and others. Metformin (Glucophage) blocks hepatic gluconeogenesis and may help to mitigate the harmful effects that excessive hyperinsulinemia has on leptin function (as we outlined above). Unfortunately, because it does block hepatic gluconeogenesis, metformin will invariably prevent a patient from re-establishing normal metabolic resiliency after they lose sufficient weight. Therefore, it is one of those reagents that is often taken off board in our patients who achieve food sobriety.

Bariatric surgery is perhaps the most definitive and effective intervention used to treat patients suffering from morbid obesity (BMI >40 kg m^{-2}) and/or patients with a BMI of 35 kg m^{-2} or greater who also suffer from metabolic syndrome. At least for gastric bypass surgery, it is felt that the anatomic rearrangements so engineered by the surgery help to decrease the influence of the gastric hormone ghrelin (which stimulates appetite at the level of the hypothalamus) and increase the output of GLP-1 and PYY, which have the opposite effect.

The Diabetic Prevention Study found metformin to be inferior to simple diet and exercise, which represent the ultimate medicinal to fight obesity and prevent T2DM. This is why weight loss means everything to the patient with T2DM or obesity and (at least for obese patients), their food sobriety demands loss of weight. In order to achieve this, many clinicians and researchers are recommending we re-examine our assumptions about the vital role that certain food stuffs play (including HFCS, EtOH in significant quantities, etc.) and the vital role that exercise plays, in stemming the tide of the obesity epidemic. The importance of considering the levels of metabolic (homeostatic), addictive (hedonistic), and relational

(symbolic) drives in patients needs to be integrated into antiobesity/food addiction research and intervention strategies.

Without a consideration of all three of these levels of intervention (metabolic, addiction, and relational), simple metabolic revisions are often trumped by addictive attractions (e.g. episodic bingeing on chocolate) or conflicted relational issues (e.g. divorce creating stress and binge eating). Therefore, it is vital that the clinician deals with the entire triad of metabolic–addictive–relational aspects of obesity. Again, this is exactly why the vast majority of weight interventions fail.

FUTURE RESEARCH

The future of research in terms of food addictions will only grow in time. First, research on evidence-based diagnostic interviewing and division into treatment-relevant subgroups is required. In the meantime, the Center for Medicaid and Medicare Services realizes the extraordinary costs of doing nothing. In terms of A-OPRs, the field is wide open. Research on epigenetic modifiers of metabolism will clarify how to more effectively triage those patients who might benefit from a specific gut peptide (such as a ghrelin blocker or GLP-1 modifier), a leptin gene-splicing intervention that ups the level of sensitivity of the leptin hypothalamic receptor, or perhaps a β-3 adipocyte agonist that more effectively uncouples energy expenditure, etc. The potential putative agents for weight loss are more or less endless as are potential genetic interventions in the future. In addition, there is a need to better research more effective bariatric surgical techniques given the likelihood that morbid obesity will only worsen over time.

There is also a need to further research our understanding of how human cognitions dealing with stress lead to increased vulnerability to food addictions. This approach needs to be cross culturally relevant. We also need to understand how to better motivate our patients to want to lose weight and afford them the supportive informational structures in order to do so. Social networking will expand in terms of importance, as will research on how to better integrate multispecialty interventions (so-called bundling) for patients suffering from food addictions.

SEE ALSO

Gambling, Exercise Dependence, Sexual Addiction, Prenatal Exposure to Alcohol and Illicit Substances, Sensory Imagery in Craving, Craving and Expectancies, Relation of Craving and Appetitive Behavior

Glossary

Acetylcholine chemical substance released from nerve endings to activate muscle, secretory glands, and other nerve cells; a key neurotransmitter.
β-3AR specific type of androgen receptor mutation.
BMI body mass index (calculated as weight in kilograms divided by the square of height in meters).
DSM-IV-TR *Diagnostic and statistical manual of mental disorders 4th edition text revision*
ErOH ethanol
GABA γ-aminobutyric acid; chief inhibitory neurotransmitter of the vertebrate central nervous system. HFCS High fructose corn syrup.
IL-6 interleukin 6; a proinflammatory marker.
MMWR *Morbidity and Mortality Weekly Report* published by the Centers for Disease Control (CDC).
NPY neuropeptide Y is a 36-amino acid peptide neurotransmitter found in the brain and autonomic nervous system.
Pulsatile beating or throbbing.
Syndrome a group of symptoms or signs that, occurring together, produce a pattern typical of a particular disease.
T2DM type 2 diabetes mellitus.
TNF tumor necrosis factor; a proinflammatory marker.
Valence the capacity of something to unite, react, or interact with something else; the degree of attraction or aversion toward/away from something.

Further Reading

Avena, N., Rada, P., Hoebel, B., 2008. Evidence for sugar addiction: behavioral and neurochemical effects of intermittent, excessive sugar intake. Neuroscience and Behavioral Reviews 32, 20–39.
Blum, K., Chen, T., Downs, B., et al., 2009. Neurogenetics of dopaminergic receptor supersensitivity in activation of brain reward circuitry and relapse: proposing 'deprivation-amplification relapse therapy' (DART). Postgraduate Medicine 121, 176–196.
Dagher, A., 2009. Overview: the neurobiology of appetite: hunger as addiction. International Journal of Obesity 33, S30–S33.
Davis, C., 2009. Psychobiological traits in the risk profile for overeating and weight gain. International Journal of Obesity 33, S49–S53.
Frazier, C., Mason, P., Zhuang, X., Beeler, A., 2008. Sucrose exposure in early life alters adult motivation and weight gain. Public Library of Science 3, e3221.
Gearhardt, A.N., Corbin, W.R., Brownell, K.D., 2009. Food addiction: an examination of the diagnostic criteria for dependence. Journal of Addiction Medicine 3, 1–7.
Gold, M.S., Graham, N.A., Cocores, J.A., Nixon, S.J., 2009. Food addiction? Journal of Addiction Medicine 3, 42–45.
Liu, Y., von Deneen, K.M., Kobeissy, F.H., Gold, M.S., 2010. Food addiction and obesity: evidence from bench to bedside. Journal of Psychoactive Drugs 42, 133–145.
Lustig, R., 2010. Fructose: metabolic, hedonic and societal parallels with alcohol. Journal of the American Dietetic Association 110, 1307–1321.
Passamonti, L., Rowe, J.B., Schwarzbauer, C., et al., 2009. Personality predicts the brain's response to viewing appetizing foods: the neural basis of a risk factor for overeating. Journal of Neuroscience 29, 43–51.
Pelchat, M., 2009. Food addiction in humans. Journal of Nutrition. 139, 620–622.
Wang, G.J., Volkow, N.D., Thanos, P.K., Fowler, J.S., 2009. Imaging of brain dopamine pathways: implications for understanding obesity. Journal of Addiction Medicine 3, 8–18.

CHAPTER

80

Gambling

Tony Toneatto

University of Toronto, Toronto, ON, Canada

OUTLINE

EPIDEMIOLOGY OF PROBLEM GAMBLING

Several epidemiological studies measuring the prevalence of problem gambling have been conducted and show lifetime prevalence rates ranging between 0.4 and 2%. With rates of pathological gambling hovering around 1% in North American surveys, the rate approaches 4% when subclinical or problem gamblers are included. Elevated rates of comorbidity have been found repeatedly in epidemiological research, as well as in studies of treatment-seeking samples. The prevalence of problem gambling appears to be lowest among older adults and highest among younger adults.

CLASSIFICATION OF PROBLEM GAMBLING

Pathological gambling was introduced into the third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III) as a "disorder of impulse control, not elsewhere classified" and comprised seven criteria predominantly related to the financial consequences of gambling (of which three needed to be met for a diagnosis), an inability to resist impulses to gambling, and negatively affected family, vocational, and personal pursuits. A differential diagnosis of antisocial personality disorder was also specified.

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00080-2

A major shift in the conceptualization of pathological gambling was reflected in DSM-III-R with a set of criteria strongly resembling substance use disorders. Nine symptoms defined the syndrome (of which four needed to be met for a diagnosis) consisting of preoccupation with gambling; gambling more than intended to; withdrawal symptoms when unable to gamble; tolerance as expressed by increasing amounts of money spent; chasing financial losses; efforts to reduce or stop gambling; disrupted social or occupational obligations; sacrifice of important social, occupational, or recreational pursuits; and continued gambling despite growing personal losses. Antisocial personality disorder, mania, and hypomania were specified as differential diagnoses.

In the fourth edition of DSM, 10 criteria defined the syndrome, of which five needed to be met for a diagnosis. As with DSM-III-R, several criteria resembled symptoms of substance dependence: preoccupation with gambling; tolerance (i.e. increasing the frequency and/or amount of money wagered); withdrawal (i.e. feeling restless or irritable when unable to gamble); difficulty stopping/reducing gambling; and sacrificing other activities (social, occupational, or leisure) as a result of gambling. The remaining criteria included chasing lost money, relying on gambling to cope with negative mood, lying about or hiding gambling, engaging in illegal activity to support gambling, and relying on others for financial relief. A manic episode was specified as a differential diagnosis. There were several differences from the previous edition of the DSM with the inclusion of gambling as a means of coping with dysphoric mood, committing illegal activities to obtain money for gambling, and relying on others to provide gambling-related financial relief in DSM-IV. Several symptoms related to the disruption of social, occupational, and personal functioning in DSM-III-R were collapsed into one criterion in DSM-IV.

As work progresses toward the fifth edition of DSM scheduled for release in 2013, the substance use disorders Workgroup of the American Psychiatric Association's DSM committee has recommended shifting pathological gambling from Impulse Disorders Not Otherwise Specified to a new category, Addiction and Related Disorders. This recommendation is based on the elevated rates of comorbidity regularly observed between substance use disorders and pathological gambling; phenomenological similarities between addiction and problem gambling; neurobiological, neurochemical and genetic commonalities; and the adoption of similar treatment approaches. In addition, the Workgroup has suggested that the name of the disorder be altered as the term "pathological" may be considered pejorative. "Disordered gambling" is being suggested as an alternative label. While DSM-IV lists 10 criteria for pathological gambling and statistical analyses identify a single underlying construct for the 10 symptoms, the criterion related to committing illegal acts (e.g. forgery, fraud) to finance gambling, which was introduced in DSM-IV, is infrequently endorsed and contributes minimally to classification accuracy. As a result, this criterion may be eliminated in DSM-V.

In addition, the Workgroup is considering modifying the cut-off score for determining a diagnosis. Currently, the cut-off for a diagnosis of pathological gambling requires endorsement of 5 of the 10 DSM-IV symptoms. However, it is not clear that this cut-off is the most valid. Although possessing face validity, it has received only limited psychometric testing. The Workgroup, based on several studies, suggests four symptoms may be sufficient to diagnose pathological gambling.

In summary, the Workgroup has suggested that listing pathological gambling among the addictive disorders section, reducing the criteria from 10 to 9, and the diagnostic cut-off from 5 to 4 should improve screening, diagnosis, treatment, and research into problem gambling.

THEORETICAL MODELS OF PROBLEM GAMBLING

Models of problem gambling have often glossed over individual differences between diverse forms of gambling, reflecting an implicit assumption that the diverse phenomenology of gambling behavior represents an unimportant variation of a more fundamental underlying process. Results of individual differences, treatment outcomes, and prevalence are typically reported in an aggregated fashion and rarely analyzed or presented by subtype of gambling. This approach to problem gambling research may present a significant barrier to a better understanding of problem gambling, its prevention, and treatment. As increasingly diverse types of gambling become available within society, the choice of game will likely be made on the basis of multiple biological, psychological, social, and cultural variables. Subtyping of problem gamblers may thus be important to increase our understanding of the assessment, cause, treatment, and natural history of problem gambling.

Choice of Gambling Activity

Given the wide availability of different forms of gambling, the choice of which games(s) to play is affected by multiple developmental, psychosocial, and demographic variables. Factors affecting the choice of

gaming may have important implications for the development of awareness, prevention, and therapeutic interventions. For example, few studies separate out slot machines and video poker even though slot machines are games of pure chance and video poker involves a degree of skill. There may be significant differences in the characteristics of those who participate in, and those who develop problems with, specific types of gambling. For example, "skill" games (e.g. track, sports) tend to be preferred by men and games that promote dissociation and narrowing of attention (e.g. electronic games, slot machines) appear to be preferred by women. In addition, games vary tremendously in terms of the speed of play, size of bet, prize structure, frequency of wins, and the role of skill.

Although each type of gambling may be conceptually distinct from another and may require different approaches to treatment, most theoretical models of problem gambling have tended to be nonspecific with respect to the type of gambling, emphasizing the putative mediating or latent variables that may underlie diverse gambling forms and generally disregarding phenomenology.

Some studies have reported differences based on gambling type. For example, gamblers on horse racing, slot machines, and mixed form gamblers treated residentially did not differ on measures of boredom proneness, depression, and sensation seeking. However, reaction times of slot machine gamblers tended to be quicker for slot machine-related words and slower for track-related words with the reverse pattern observed for racing gamblers. In a comprehensive study, treatment-seeking gamblers were compared on the basis of their preferred gambling activity (i.e. sports, racing, cards, slots, lotteries) differed in South Oaks Gambling Screen (SOGS) scores (i.e. racing gamblers the highest), gambling frequency (i.e. lottery players the highest), money wagered (i.e. racing gamblers the highest), years of problem gambling (i.e. racing gamblers the highest), and Gamblers Anonymous (GA) participation (i.e. least likely by lottery players) as well as in patterns of psychiatric illness and addiction. Horse/dog race gamblers tended to be less educated, men, and older with an earlier onset of gambling. Demographic variables also distinguished the gambling subtypes. Sports gamblers tend to be younger men with relatively higher rates of addiction comorbidity. Electronic gaming gamblers (e.g. slot machines) were more likely to be female and older with higher rates of psychiatric comorbidity and later onset of gambling. The lottery gamblers tended to have both alcohol and other psychiatric comorbidity and the highest frequencies of gambling.

Gender appears to affect the choice of gambling activity. Women preferred nonskill/nonstrategic forms of gambling such as bingo and slot machines, whereas men preferred skill/strategic types of gambling such as sports and track gambling. Gambling choice may also reflect gender differences in motivation for gambling (i.e. emotional dampening versus excitation).

Many have suggested that there are at least two main subtypes of gamblers: (1) overstimulated gamblers, who are attracted to types of gambling that may soothe or diminish dysphoric emotional states by engaging in games that promote dissociation or narrowing of attention (e.g. slot machines, electronic gaming) and (2) understimulated gamblers, who are drawn to games that provide stimulation and arousal (e.g. sport betting, race track betting, craps). These two models have been incorporated into the Pathways Model, a synthesis of the available knowledge on the determinants and variables associated with developing problem gambling, including the data on the biological correlates of gambling. Three main pathways that lead to problem gambling have been distinguished: behaviorally conditioned (BC), emotionally vulnerable (EV) (i.e. overstimulated), and antisocial impulsivist (AI) (i.e. understimulated) gamblers. The Pathways Model assumes all gamblers, independent of pathway, will be influenced by gambling availability, operant (i.e. schedules of reinforcement) and classic (i.e. stimulus control) conditioning, and irrational beliefs.

The EV gambler generally demonstrates higher rates of mood and anxiety symptoms, proneness to abuse psychoactive substances, relies on gambling to cope with negative mood, poor problem-solving skills, higher levels of psychological distress, fewer legal problems, and dysfunctional family history. Women tend to have a greater representation among the EV subtype, consistent with studies that have found affective disorders and escape gambling more prominent among female problem gamblers.

The BC gamblers are sensitive to irrational and distorted beliefs about gambling, show evidence of poor decision making, often lack any concurrent psychiatric comorbidity, do not have a familial psychiatric history, and tend to show lower severity of gambling problems. BC gamblers may often be underrepresented in treatment-seeking populations.

The AI gambler tends to show the highest severity of gambling and substantial psychopathology, have lower educational attainment, are predominately male, and show evidence of neurochemical and neurological disturbance. Character dysfunctions are common in this subtype, especially those associated with antisocial personality, impulsivity and attention deficits, addictive behavior, suicidality, criminality, and low frustration tolerance.

SCREENING AND DIAGNOSING PROBLEM GAMBLING

Accurate identification of disordered gambling is aided by the growing development of assessment instruments specific to gambling pathology. Available tools establish the severity of problem gambling (i.e. degree of involvement, consequences associated with gaming behavior) with the goal of classifying the gambler along a continuum of severity and/or making a diagnosis. Many of the same measures are also used as screening devices to distinguish problem gamblers from nonproblem gamblers.

Screening Measures

One of the earliest instruments to screen for problem gambling, the Twenty Questions (GA-20Q), was developed within GA. Evidence for the reliability and concurrent, convergent, and predictive validity data for the GA-20Q in three independent samples involving 456 participants (two samples of problem gamblers in treatment and a nontreatment sample of problem gamblers) has been established. Classification analyses indicated that the 20Q is comparable with the DSM-IV diagnostic criteria for pathological gambling in specificity, sensitivity, and rates of false-negatives and false-positives.

The SOGS was developed in the mid-1980s and is arguably the most-well known self-report screening tool for gambling pathology. The use of the SOGS has extended from a screening instrument to nonclinical populations and prevalence surveys, raising questions about the psychometric properties and classification accuracy of the SOGS under these conditions. Furthermore, development of the SOGS items was based on the DSM-III and DSM-III-R criteria, and do not reflect the evolution of criteria observed in DSM-IV. Attempting to address concerns with the SOGS and develop a more appropriate measure of problem gambling for use in the general population, the Canadian Problem Gambling Index (CPGI) offers a more "holistic view of gambling" within a social context. The CPGI consists of three main sections: gambling involvement, problem gambling assessment, and correlates of problem gambling (including familial history of gambling); and yields five categories of gambling behavior (ranging from nongambling to problem gambling). Initial studies indicate that the CPGI demonstrates good reliability and validity.

For clinicians and researchers interested in a very brief self-report instrument using DSM-IV criteria, the Lie/Bet questionnaire contains only two items (e.g. need to bet more money, lie about gambling). A three-item version of the National Opinion Research Center DSM Screen for Gambling Problems (NODS), consisting of symptoms related to loss of control, lying, and preoccupation (the "CLiP"), and requiring 1 min to administer has been developed. The CLiP has excellent sensitivity and specificity, and identified almost all pathological gamblers and problem gamblers diagnosed by the full NODS. Recently, a three-item measure, the Brief Biosocial Screen (BBS), has been developed to detect gambling disorders among gamblers in the general household population and treatment seekers. The three symptoms reflect the neuroadaptive aspect (i.e. withdrawal), psychosocial characteristics (i.e. lying), and social consequences (i.e. obtaining money to gamble from others) of gambling. The BBS has been shown to have high sensitivity and high specificity.

Diagnostic Measures

Currently there is no well-established, structured, clinical interview for assessing problem gambling. However, the Diagnostic Interview for Gambling Severity (DIGS) and the Diagnostic Interview Schedule (DIS) Pathological Gambling Module (GAM-IV) have encouraging reliability and validity. Excellent sensitivity and specificity of the DIGS in classifying pathological gamblers in the general population has been reported.

The NODS, a DSM-IV–based telephone screening measure to identify problem gambling has been assessed as a potential outcome measure for gambling treatment studies. Internal reliability was fair to good. The factor structure revealed three subfactors: measuring negative behavioral consequences, preoccupation/impaired control over gambling and tolerance, and withdrawal/relief gambling. The NODS total score correlates moderately with gambling behavior and outcome and the SOGS. The NODS displayed good reliability, concurrent and discriminant validity, and may be a clinically useful screen for gambling problems among patients with substance use disorders.

Practical Issues in Selecting Instruments

There is considerable variability in screening and assessment instruments ranging from as few as two items and only minutes to administer to extensive interviews requiring over an hour to administer. Some instruments can be self-administered; others must be administered by clinicians. Although some measures only indicate the severity of a gambling problem, some measures yield a diagnosis. The SOGS tends to yield higher false-positive rates. DSM-based measures may produce higher false-negative rates, missing individuals with milder gambling problems. The SOGS is ideal when the goal is to identify possible gambling problems; the NODS is to be indicated if the goal is to derive a diagnosis of pathological gambling.

RISK FACTORS AND PROBLEM GAMBLING

Recent efforts to increase our understanding of the psychology of problem gambling has focused on a number of risk factors, including the role of negative affect (depression and anxiety), personality factors (sensation seeking, impulsivity, boredom proneness, extroversion, locus of control, narcissism, and antisociality), concurrent disorders (overlap between gambling pathology and mood disorders, substance misuse, and attention-deficit hyperactivity disorder), gender differences, and the role of cognition (including metacognition).

Very few risk factors for problem gambling have been well established (i.e. supported by more than two studies). Some risk factors with supportive evidence include demographic variables (age, gender), cognitive distortions (erroneous perceptions, illusion of control), sensory characteristics, schedules of reinforcement, and psychiatric comorbidity (obsessive–compulsive disorder, drug abuse), and criminal activity. Relevant research is highlighted below.

Demographic Variables

Age

Age and gambling disorders have consistently been shown to be negatively correlated. Younger people tend to show higher rates of problem gambling than older adults. The earlier the onset of gambling, the greater the probability of developing a more severe gambling disorder. In a longitudinal study of males of low socioeconomic status (SES) beginning at age 11 years, three groups were identified: an early-onset/high gambling involvement, late-onset/high gambling involvement, and a low gambling group. Anxiety and impulsivity during childhood and early adolescence distinguished the early-onset and the nonproblem groups, with the late-onset group falling in between. Gender differences in gambling preferences have been found among adolescents, with boys preferring skill games (e.g. cards) and girls preferring lotteries. Older adults may prefer games that are less competitive and demand fewer cognitive and attentional resources such as bingo, lotteries, and slot machines.

Gender

Men have consistently been shown to comprise a higher proportion of problem gamblers in most jurisdictions throughout the world. An important gender difference has been the "telescoping effect": the observation that women appear to experience a different illness course than men. The interval between the onset of gambling and its recognition as a problem appears to be shorter for women. Gender differences in motivation have also been reported. Women's motivation to gamble was more likely to be associated with feelings of loneliness and dysphoria and escape from personal or family problems, whereas men gambled more for action, excitement, and arousal, and to win money. Electronic gaming machine gambling, in particular, may induce dissociative states that women may rely on to escape from other life problems to a greater degree than men. In a study of treatment-seeking gamblers, a higher proportion of the electronic gaming machines and bingo gamblers were women and a higher proportion of the horse racing, legalized betting, and card gamblers were men.

SES

Lower SES, unemployment, low educational achievement, and low income have been consistently associated with higher rates of problem gambling. Minority status and SES were found to be significantly associated with gambling problems after controlling for gambling behavior, addictive behavior, and other sociodemographic variables. SES may also interact with type of gambling; that is, electronic gaming machine (e.g. slot machines) players tended to be of lower SES than roulette players.

Marital Status

Problem gamblers are more likely to be divorced than nonproblem gamblers. Lower rates of problem gambling among married individuals compared with the divorced or separated have been shown in Swedish, Norwegian, and Australian samples. There is also evidence that married problem gamblers may be more likely to seek treatment.

Psychiatric Disorders

Mood

The association of pathological gambling with depressive mood is one of the most studied risk factors for pathological gambling. Investigations in this area have produced mixed results, with some research initiatives unable to establish a connection between depression and gambling. In contrast, several studies have highlighted the high rates of co-occurrence of gambling and depressive symptoms. Epidemiological studies have reported up to 50% of problem gamblers may have a lifetime mood disorder.

Studies that have assessed directionality between mood and gambling find that the gambling disorder typically precedes the onset of depression. Affective symptoms were equally likely to either precede or follow the onset of a gambling problem in a sample of treated pathological gamblers.

Anxiety

Epidemiological studies have reported up to 40% of problem gamblers may have a lifetime anxiety disorder, with even higher rates among treatment-seeking gamblers (Ibanez et al., 2001). Older samples tend to have high rates of lifetime anxiety disorder (47.5%), obsessive–compulsive disorder (37.5%), and panic disorder (27.5%). Similarities between problem gambling and obsessive compulsiveness have been specifically noted, particularly the preoccupation with gambling and the amount of time spent developing gambling strategies. Obsessive–compulsive factors in lottery and scratch ticket gamblers also showed that pathological gamblers endorsed more symptoms of obsessive compulsiveness compared with recreational gamblers.

Substance Use

Comorbidity between gambling pathology and substance use disorders has often been observed. The majority of pathological gamblers in a national sample of 43 000 respondents had a lifetime alcohol use disorder and 38% had a lifetime drug use disorder. Substance abusers with a gambling problem reported increased levels of somatization, obsessive compulsiveness, interpersonal sensitivity, and paranoia. Problem gamblers with a history of treatment for substance abuse reported more depression, hallucinations, suicidal ideation and attempts, and difficulty controlling violent behavior over their lifetime versus gamblers with no previous treatment for substance abuse.

The relationship between lifetime diagnoses of drug dependence or mood disorder and gambling treatment outcomes in a sample of quit (either with or without formal treatment) gamblers followed up for up to 5 years was studied. Those with lifetime drug diagnoses were less likely to have established a period of abstinence for more than 3 months in duration. Those with a lifetime history of a mood disorder also took longer to establish 3 months of abstinence. Comorbid addictive and mood disorders predicted shorter-term, but not longer-term, gambling abstinence. These results illustrate the importance of assessing lifetime comorbid disorders, which clearly complicate the recovery process for pathological gambling.

Individual Differences

Sensation Seeking

Individuals who are underaroused, prone to boredom, hypomanic, or depressed may value the stimulation and excitement associated with waging higher amounts of money and participating in higher-risk gambling activities; this has often been proposed as an important individual difference between problem gamblers. Sensation seeking has been among the most commonly investigated personality factors among problem gamblers but the results have been unequivocal. Correlation between higher needs for arousal and stimulation and preference for riskier bets has been observed in undergraduate gambling. In a sample of community-recruited problem gamblers, some aspects of sensation seeking (i.e. the boredom susceptibility) were found to correlate significantly with severity of gambling but negatively correlated in a study of pathological gamblers. Horse race gamblers (a "skill" game) sought heightened arousal but electronic gaming machine players (a chance game) appeared to seek dampened arousal. Among youth, sensation seeking was associated with gambling frequency but not with the development of a gambling disorder.

Physiological Arousal

Physiological "rush" or excitement when gambling has often been proposed as a major reinforcing stimulus for gambling. This appears to be supported by research with sports, horse race and dog race gamblers who showed increases in blood pressure and heart rate while gambling. Many studies have demonstrated increased levels of arousal in problem gamblers, especially in ecologically valid settings. In the latter studies, the effects were enhanced when the gamblers played in a real casino versus a laboratory setting. However, not all studies have found a relationship between arousal and problem gambling or even between objective and subjective reports of arousal. Winning during slot machine play was associated with an increased heart rate. An increase in heart rate during electronic machine gambling was observed but only when winning. The positive correlation between winning and heart rate did not require that the gamblers actually needed to ultimately win money.

The inconsistent results in the arousal/sensation-seeking literature, similar to other constructs discussed in this chapter, can be accounted for by a multitude of methodological differences (i.e. measure of arousal, type of gambling, setting, psychometric instruments, demographics, gender). As noted elsewhere, the inclusion of heterogeneous samples of gamblers often introduces variability in the results.

Cognitive Variables

A substantial body of work is devoted to the role of cognitive factors in pathological gambling. Two major beliefs seem to describe the problem gambler's irrational cognition. *Primary illusory c*ontrol, the irrational belief that the gambler can control the outcomes of gambling events, has been distinguished from secondary illusory control, the irrational belief that the gambler can predict the outcomes of gambling events.

Even games that are ostensibly completely random such as slot machines can elicit irrational beliefs about control and prediction. These core beliefs yield a wide array of irrational or maladaptive beliefs about gambling outcomes that have been well described in the literature.

The raw frequency of erroneous perceptions may not distinguish problem gamblers from nonproblem gamblers, at least in a sample of video lottery players. Pathological gamblers were more convinced of the truth of their erroneous perceptions than the nonproblem gamblers. Although the nonproblem gamblers demonstrated a reduction of confidence in their erroneous perceptions as their play continued, the pathological gamblers showed a trend toward increasing confidence over time. These results suggest a role for a metacognitive level of analysis of gambling cognition.

Educating university students on probability theory (e.g. odds) through the use of gambling examples produced differences in the ability to calculate gambling odds and resistance to irrational gambling-related mathematical beliefs compared with those who were instructed on probability theory generically (i.e. without the aid of gambling-related examples). Surprisingly, however, there was no effect on gambling behavior. Mathematical knowledge related to gambling did not translate into a modification of gambling behavior. In another study, individuals who were reminded about the independence of events reported a decreased motivation to pursue playing roulette and made fewer erroneous perceptions.

Early Childhood Experiences

Rates of abuse and neglect among problem gamblers are higher than among nonproblem gamblers. Almost two-thirds of gamblers attending a residential treatment facility reported a history of emotional trauma, 40% reported a history of physical trauma, and 24% a history of sexual trauma, mostly during childhood. All females reported some form of abuse compared with 61% of the males. A positive history of childhood trauma was associated with suicidal, addictive, and psychiatric symptoms.

Influence of Parental Gambling

Adult problem gamblers have reported higher rates of gambling behavior by their parents compared with nonproblem gamblers. In a study of high school students, probable pathological gamblers reported higher likelihood of parental gambling, suicide attempts, alcohol and drug misuse, and peer pressure susceptibility. The transmissibility of gambling-related irrational beliefs in nonproblematic video lottery play was demonstrated in a study that found the father's gambling-related cognitions and behaviors were most strongly correlated with their offspring's gambling behavior. In another study, subjects took greater monetary risks while playing video lotteries in the presence of an accomplice who had expressed erroneous cognitions about gambling but not in the presence of an accomplice who had expressed rational beliefs. These results may help explain how childhood exposure to gambling environments affect gambling behavior later in life.

PSYCHOLOGICAL TREATMENT OF PROBLEM GAMBLING

The psychological treatments that have been evaluated in controlled trials can generally be classified as brief and cognitive-behavioral in orientation. The focus has generally been placed on directly modifying gambling-related behaviors and cognitions, and in developing behavioral and cognitive coping skills to reduce gambling frequency, gambling expenditures, or urges using problem-solving, practical, and pragmatic approaches. Several reviews of the controlled treatment literature have shown cognitive-behavioral treatment (CBT) approaches to be supported with the best empirical evidence, although this finding generally reflects the fact that non-CBT studies tend to be evaluated less frequently. The better-designed cognitive-behavioral studies have obtained long-term improvements in gambling of between 50 and 75%.

CBT

Cognitive-behavioral techniques to treat pathological gambling have included alternative activity planning, problem solving, limit setting, social skills, communication training, relapse prevention, stimulus control, and in vivo exposure. Individually administered response-prevention treatment was superior to group cognitive therapy or the combined treatment at the 1-year follow-up with no group differences between the active treatments for any of the other dependent variables (e.g. frequency of gambling). The addition of cognitive therapy was not found to be differentially more effective than behavioral treatment or a combined CBT.

Slot machine gamblers who were first treated with stimulus control and in vivo exposure followed by response prevention until abstinent was achieved were subsequently randomly assigned to one of three relapse prevention (RP) modalities: individual RP, group RP, and no-treatment control. Significant differences in abstinence rates between the two RP groups (which did not differ) and the no-treatment control group

emerged at 3 months after treatment and were maintained throughout the 12-month follow-up.

Due to high attrition, compliance enhancements (e.g. follow-up letters, positive reinforcement, increased self-efficacy) were added to a cognitive-behavioral intervention. Subjects who received the compliance enhancement completed the eight-session program at a higher rate (65%) compared with those randomly assigned to the standard CBT (35%) but there were no group differences on measures of improvement (e.g. SOGS, % net monthly income gambled) or at the 9-month follow-up.

Node-link mapping-enhanced cognitive-behavioral therapy (CBGT mapping) was compared with 12-step facilitated (TSF) treatment. A wait-list control group was also included. The group treatments resulted in significant improvements on the dependent measures, while the wait-list group showed little change. Improvements were found on DSM-IV criteria, self-efficacy, and gambling episodes. Both intervention groups sustained treatment gains at the 6-month assessment but demonstrated no significant differences in terms of gambling outcomes. Relatively few TSF members attended GA, indicating that attending GA meetings may not be necessary for treatment to be effective.

Eight sessions of CBT or the 12-step treatment-oriented approach (based on the first five steps of GA) were compared in a quasiexperimental research design. At 12 months after treatment, there were no group differences on key gambling variables (e.g. frequency, abstinence rates, money wagered) in an analysis of completers, with both treatments significantly reducing the frequency and amount of money spent on gambling. Participants who attended more sessions and chose an initial abstinent treatment goal seemed to achieve better outcomes.

In one of the largest evaluations of CBT, pathological gamblers were randomized to GA, GA referral plus a self-directed CBT manual (eight chapters), or GA referral plus individual CBT (eight sessions) with a 12-month follow-up. Primary outcome variables were symptom severity scores, days of gambling, and money gambled. Patients in the GA referral plus individual CBT group had significantly greater improvement than patients in the GA referral only group on all primary outcome measures and greater longer-term benefits relative to the other two groups.

An integrated 14-week treatment for comorbid problem gambling, anger, and substance use randomized to either treatment of anger and addictions or treatment-as-usual (TAU) for gambling and substance use showed that the integrated anger and addiction treatment showed significantly less gambling after treatment and at follow-up and less trait anger and substance use at the 1-year follow-up.

A sample of inpatient pathological gamblers treated with CBT found over half of the sample abstinent at the 1-year follow-up and an additional 16.3% were abstinent until the last 3 months of the follow-up. However, the low follow-up response rate of just over half of the initial sample (i.e. 52%) limits the generalizability of the study. Patients who relapsed reported higher psychological distress at the end of treatment and at follow-up and lower quality of life at follow-up.

Cognitive Treatments

Although treatments targeted at correcting cognitive distortions are effective, it is too early to tell whether treating cognitive distortions is a prerequisite for successful gambling treatment outcomes. Research on cognitive pathology specific to different types of gambling is underway (e.g. slot machines; lottery players), but much more is necessary before the cognitive characteristics of different types of gambling and their role in the course of problem gambling is understood.

Several randomized control studies attest to the efficacy of cognitive treatment for problem gambling. Cognitive intervention targeting erroneous beliefs found significant reductions in gambling severity, frequency of gambling, and amount wagered compared with the wait-list control group in several studies with results maintained at a 12-month follow-up in some studies and up to 2 years in others.

Although cognitive distortions among problem gamblers have been well established, less is known about whether treating cognitive distortions is more effective than other therapeutic approaches to gambling. Four brief, six-session, homogeneous treatments for problem gambling (i.e. cognitive, behavioral, motivational, minimal intervention) were compared in a sample consisting primarily of middle-aged, unmarried, underemployed men recruited from the community. The results demonstrated that a cognitive approach was not superior to the three other treatments, which did not explicitly address cognitive distortions. There are likely multiple pathways to therapeutic change that may not necessarily require the modification of gambling-related cognitive psychopathology.

Brief and Motivational Interventions

In recent years brief treatments for problem gambling emphasizing brief motivational interventions (BMI) have emerged. A unique aspect of BMI is the willingness to acknowledge and accept ambivalence about modifying gambling behavior. A firm commitment for change is not insisted upon, and participants are encouraged to discuss the pros and cons of gambling. BMI encourages the exploration of the values and principles that govern

their life without demanding change or assigning judgment.

Combining BMI with CBT may help reduce dropout rates. A brief minimal intervention for gamblers compared a self-help manual with the manual plus telephone/postal contact. The manual-only group reduced the number of gambling sessions per week and dollars wagered per week for the 6 months after receiving the manual, whereas the manual/interview subjects showed a reduction between the interview and 3 months only.

Two versions of a self-help manual based on CBT (i.e. mailing the self-help manual versus preceding the mailing of the manual with a telephone BMI reviewing assessment information and enhancing commitment to change) were compared with a wait-list control group. Significant reductions in gambling behavior were reported by 84% of subjects over the 1-year follow-up period. An initial superior outcome demonstrated by the BMI plus manual group over the manual-only group was evident at the 3-month follow-up but not at 12 months, except for gamblers with less severe gambling.

The efficacy of a single BMI session compared with a no-treatment control interview showed that the BMI condition significantly reduced gambling expenditure and gambling days per month and led to less emotional distress than the control condition at the 1-year follow-up. However, problem gambling severity decreased (as measured by the SOGS and CPGI) in both conditions. The provision of the self-help manual, which the no-treatment control received, may have reduced between-group differences.

A comparison of four groups of non–treatment-seeking gamblers (no-treatment control, 10 min of brief personal feedback and gambling advice, a 50-min motivational enhancement session, or one session of motivational enhancement plus three sessions of CBT) showed that those in the brief advice condition showed significantly decreased gambling compared with the no-treatment control group during the 6-week study period. No additional gains were made in the remaining 9 months of the follow-up period. The motivational enhancement group did not differ from the control group. Those in the motivational enhancement plus CBT group were no different from the control group during the 6-week study period, but were significantly better on some outcome measures throughout the 9-month follow-up period.

A comparison of four sessions of motivational interviewing (MI) and eight sessions of CBT with a wait-list control found both treatments were superior to the control group on a measure of severity (i.e. NODS) and depression but on no other outcome measure. There were no differences between MI and CBT post-treatment or throughout the follow-up period although both treatments showed significant within-treatment gains. High rates of attrition (mean of 5.6 out of eight CBT sessions attended; mean of 0.9 sessions out of four MI sessions attended) compromised the generalizability of the findings. In this study, the no-treatment control group made significant gains on the basis of non–treatment-related factors (i.e. due possibly to intrinsic motivation natural recovery).

Subtyping and Treatment Outcome

Treatment-seeking pathological gamblers were classified according to the subtypes described in the Pathways Model and evaluated as to whether there would be a differential benefit from CBT or GA referral through a 12-month follow-up. Compared with BC gamblers, EV gamblers had higher baseline psychiatric and gambling severity, and were more likely to have a parent with a psychiatric history. AI and EV gamblers experienced greater gambling severity throughout treatment than BC gamblers, but all three subtypes demonstrated similar positive responses to treatment and improved at similar rates. As a result, the EV and AI subtypes demonstrated greater problem gambling severity post-treatment and during the follow-up period relative to BC gamblers due to higher gambling severity before treatment. Subtyping problem gamblers according to the Pathways Model may not be useful for making specific treatment recommendations.

Meta-analysis of CBT

Meta-analyses of CBT or variants show that three forms of CBT (generic CBT, MI, or combined MI/CBT) demonstrated large and significant effect sizes through follow-up periods of up to 2 years, suggesting the enduring effects of CBT despite variability in populations being treated, recruitment methods, gambling severity, treatment type, outcome measures, and study design. Although this result lends support to the use of CBT, it cannot be used to claim that other treatments would be ineffective. Few studies directly compare two different treatment modalities.

Although the heterogeneity of gambling problems may preclude a single therapeutic approach to CBT, this modality was effective regardless of the type of problem gambling at both the 3- and 6-month time windows. This indicates that there is a common mechanism underlying gambling behavior that can be targeted by CBT regardless of how the gambling behaviors are phenomenologically expressed. An alternative pathway for the clinical efficacy of CBT is that gambling behaviors are not directly affected by CBT at all. Comorbid psychiatric syndromes (e.g. mood, anxiety) may improve first

with a concomitant reduction in gambling behaviors as a consequence. Metaregression effects suggest that treatment outcome is independent of the therapy length. This is encouraging for brief therapeutic programs as several studies have demonstrated that a one-session, motivational intervention could perform as well as longer (i.e. six sessions) CBT.

A closer study of efficacious treatments for problem gambling reveals that they tend to be multimodal, combining cognitive, behavioral, and motivational interventions. Thus, it is difficult to evaluate which class of interventions is the most effective in effecting clinical change. This can be a serious weakness in the formulation of optimally effective treatment for problem gambling given the brief duration of contact between problem gamblers and the treatment system and high attrition rates among gamblers in treatment. To maximize the impact of treatment and reduce the risk of noncompliance, dropout, and relapse, treatments should strive to include only those interventions that have been empirically supported.

PHARMACOLOGICAL TREATMENTS OF PROBLEM GAMBLING

A wide variety of medications have been evaluated for their efficacy in the treatment of problem gambling. Two classes of medications have received considerable experimental investigation: opioid antagonists and antidepressants.

Opioid Antagonists

Two opioid antagonists, naltrexone and nalmefene, have been studied in several double-blind, placebo-controlled, randomized trials. Naltrexone-treated patients were more likely to improve (75%) compared with patients in the placebo condition (24%). Effects on actual gambling behavior were not reported in either study. Post hoc analyses in one study further support the hypothesis that pharmacological manipulation of the opiate system may target core symptoms of pathological gambling. However, naltrexone was not superior to placebo in reducing alcohol and gambling behavior in a concurrently comorbid population of alcohol-dependent/abusing pathological gamblers.

Antidepressants

Several antidepressants have also been evaluated with mixed results. Fluvoxamine and paroxetine have been shown to reduce gambling expenditure and time spent gambling in double-blind placebo-controlled studies. As no data on actual gambling behavior were collected, it is not known whether paroxetine actually modified gambling behavior. Fluvoxamine was compared with topiramate, an anticonvulsant that targets gamma-aminobutyric acid (GABA)ergic and antiglutamatergic mechanisms, in a randomized, blind-rater comparison study. Rates of abstinence were equal in both groups (75%) but the rate of attrition was much higher in the fluvoxamine group (50%) compared with 20% in the topiramate group.

Based on the available research, there is no compelling empirical evidence for the efficacy of any medication except for naltrexone. To date, no medication has been approved for pathological gambling by the US Food and Drug Administration.

SUMMARY OF TREATMENT EFFICACY

Review of the best-designed treatment studies indicate that CBT and pharmacological treatments for gambling can be judged as possibly efficacious. No specific modality or intervention has been found to be unequivocally effective by at least two independent research teams. The significant findings for CBT have usually been reported in comparison with wait-list controls and less often against other credible treatments. CBT, even when delivered via a manual and involving only minimal therapist contact, has the most empirical support compared with no treatment. Among the medications that have been studied, there is some limited support for naltrexone. However, it is not possible to determine which specific type of CBT or medication is most effective or whether CBT is more effective than other treatment options.

The issue of attrition remains an important methodological issue to consider in evaluating treatment efficacy, given its common occurrence. Among gambling treatment studies published between 1982 and 2005, rates of attrition in pharmacological treatment studies ranged from 11 to 40% (in short-term studies) and from 48 to 59% (in long-term studies). Between one-third and one-half of subjects in psychosocial treatment studies dropped out of treatment; the rates of attrition in community multimodal treatments ranged between 29 and 83%. GA treatment studies showed similar rates of attrition (i.e. 50–69%). Although there is evidence that more treatment is generally associated with better outcomes, some recent work has challenged this finding.

Placebo/nonspecific response to treatment by problem gamblers has often been identified as a factor common to both pharmacological and psychosocial treatments. High rates of placebo response rate (up to 72%) have been reported. The average rate of placebo response among pathological gamblers is estimated to be ~40%. Observations that one-third of early treatment

dropouts and treatment referrals who did not engage in treatment were abstinent from gambling at the 6-month follow-up suggest that placebo or nonspecific response can account for the benefits of treatment. Nonblinded studies with nonequivalent control groups are more likely to overestimate treatment effects by including nonspecific effects with the specific benefits of treatment.

Implications for Delivery of Treatment

With the prevalence of concurrent addiction and psychopathology common among gamblers well established, appropriate screening for such comorbidity and determination of the primary disorder should influence decisions about appropriate treatment (most likely a combination of psychological and pharmacological therapies). Very little research has investigated the outcomes following the combination of validated psychological therapies and medications.

For gamblers who do not have a concurrent disorder, the initial intervention should strive to increase the individual's commitment to treatment and resolve treatment-disrupting ambivalence as much as possible through BMI. Motivational treatments may serve this purpose and thus reduce the relatively high rates of dropout found in most gambling treatment studies. The high attrition rates suggest that more effort must be made to strengthen the client's motivation to change.

With strengthened motivation, the available empirical research suggests that CBT, generally brief and delivered on an outpatient basis, seems to be the most effective treatment. Concurrent pharmacotherapy (i.e. naltrexone, selective serotonin reuptake inhibitors) remains promising as an adjunct to psychosocial treatment, but is insufficiently empirically validated at the present time.

SEE ALSO

Internet Addiction: Cybersex, Video Game Addiction, Cognitive Factors in Addictive Processes, Natural Recovery, Craving and Expectancies, Models of Relationships between Substance Use and Mental Disorders

List of Abbreviations

AI antisocial impulsivist
BBS Brief Biosocial Screen
BC behaviorally conditioned
BMI brief motivational interventions
CBT cognitive-behavioral treatment
CLiP control, lying, and preoccupation
CPGI Canadian Problem Gambling Index
DIGS Diagnostic Interview for Gambling Severity
DSM *Diagnostic and Statistical Manual of Mental Disorders*
EV emotionally vulnerable
GA Gamblers Anonymous
NODS National Opinion Research Center DSM Screen for Gambling Problems
RP relapse prevention
SES socioeconomic status
SOGS South Oaks Gambling Screen
TSF 12-step facilitated

Further Reading

Blaszczynski, A., Nower, L., 2002. A pathways model of problem and pathological gambling. Addiction 97, 487–499.

Gooding, P., Tarrier, N., 2009. A systematic review and meta-analysis of cognitive-behavioural interventions to reduce problem gambling: hedging our bets? Behaviour Research and Therapy 47, 592–607.

Hodgins, D.C., Currie, S.R., Currie, G., Fick, G.H., 2009. Randomized trial of brief motivational treatments for pathological gamblers: more is not necessarily better. Journal of Consulting and Clinical Psychology 77, 950–960.

Kessler, R.C., Hwang, I., LaBrie, R., et al., 2008. DSM-IV pathological gambling in the National Comorbidity Survey Replication. Psychological Medicine 38, 1351–1360.

Milosevic, A., Ledgerwood, D.M., 2010. The subtyping of pathological gambling: a comprehensive review. Clinical Psychology Review 30, 988–998.

Milton, S., Crino, R., Hunt, C., Prosser, E., 2002. The effect of compliance-improving interventions on the cognitive-behavioral treatment of pathological gambling. Journal of Gambling Studies 18, 207–229.

Petry, N.M., 2005. Pathological Gambling: Etiology, Comorbidity and Treatment. American Psychological Association, Washington, DC.

Petry, N.M., Ammerman, Y., Bohl, J., et al., 2006. Cognitive–behavioral therapy for pathological gamblers. Journal of Consulting and Clinical Psychology 74, 555–567.

Petry, N.M., Weinstock, J., Ledgerwood, D., Morasco, B., 2008. A randomized trial of brief interventions for problem and pathological gamblers. Journal of Consulting and Clinical Psychology 76, 318–328.

Petry, N.M., 2010. Pathological gambling and the DSM-V. International Gambling Studies 10, 113–115.

Shaffer, H.J., Hall, M.N., 2001. Updating and refining prevalence estimates of disordered gambling and behavior in the United States and Canada. Canadian Journal of Public Health 92, 168–172.

Smith, D., Hodgins, R., Williams, R. (Eds.), 2007. Research and Measurement Issues in Gambling Studies. Elsevier, San Diego, CA.

Toneatto, T., Ladouceur, R., 2003. The treatment of pathological gambling: a critical review of the literature. Psychology of Addictive Behaviors, 284–292.

Toneatto, T., Nguyen, L., 2007. Individual characteristics and problem gambling behavior. In: Smith, G., Hodgins, D., Williams, R. (Eds.), Research and Measurement Issues in Gambling Studies. Elsevier, San Diego, CA, pp. 280–304.

Westphal, J., 2008. How well are we helping problem gamblers? An update to the evidence base supporting problem gambling treatment. International Journal of Mental Health and Addiction 6, 249–264.

CHAPTER

81

Internet Addiction: Cybersex

Qiaolei Jiang[*], *Xiuqin Huang*[$], *Ran Tao*[$]

[*]The Chinese University of Hong Kong, N.T., Hong Kong SAR, China

[$]General Hospital of Beijing Military Region, Beijing, China

OUTLINE

DEFINING CYBERSEX

As Internet has become a popular medium all over the world, more and more individuals are engaging in various kinds of sexual exchanges online. Cybersex, Internet sex, net sex, or computer sex refers to an enormous range of real-time sexual encounters by using computer-mediated interactive technologies. This rapidly growing phenomenon has drawn much attention due to its novelty since the 1990s.

Cybersex can be defined as a subcategory of online sexual activity where Internet is used for sexually gratifying activities. In the computer-mediated communication environment, cybersex participants engage in various sexual experiences, and Internet has become a virtual playground for sexual encounters. Cybersex is a kind of interactive erotic experience, which usually involves two or more participants having sexual exchanges online with the purpose of sexual arousal and stimulation.

Participants connected remotely via a computer network interact with one another through sexually explicit messages, such as typed text, live video, and various digitized sexual content, written, auditory, or visual. Cybersex is commonly performed in Internet chat rooms and on instant messaging systems. Newsgroups and online role-playing games can also be the starting points for cybersex. Individuals engage in cybersex in many ways but basically, two forms are most prevalent: text cybersex and televideo cybersex. Cybersex users engage in simulated sex talk while online for the purposes of sexual pleasure and may or may not include masturbation. Cybersex can either be a goal in itself or may serve as a first step toward a real-life encounter. Related activities include phone sex with people met online, and online affairs that progress to skin-to-skin sex.

TEXT CYBERSEX

Text cybersex refers to a form of role-playing in which the participants pretend they are having actual sexual relations through real-time written communication of

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00081-4

erotic-sexual content. The conversations vary, ranging from flirting and talking dirty, to very detailed descriptions of having intercourse.

In text cybersex, participants try to get mutual stimulation by describing their sexual fantasies in detail and responding to their chat partners. Through online chat, people type out their appearances, erotic actions, utterances, touches, feelings, and happenings to one another. In this way, text cybersex participants make online chat to be a personalized interactive arena for sexual experience by provoking, constructing, and playing out sexual encounters.

To a certain degree, the online "hot" chats or sex talk can be seen as extensions of telephone sex and party lines. Similar to telephone sex, text-based cybersex is purely communicative, without co-present bodies and actions. While different from phone sex, there can even be no spoken words. Thus, text cybersex is a form of co-authored erotica interactively crafted by anonymous participants in a process of communication.

The use of language in text cybersex is often creative, since the whole experience depends on the exchange of detailed and intimate typed conversations. The quality of a cybersex encounter typically depends on the participants' abilities to evoke a vivid, visceral mental picture in the minds of their partners. Text cybersex participants usually draw from a vast repertoire to describe the enormous range of bodily sensations which typically accompany sex, such as appearances, expressions, gestures, odors, and physical sensations, which may require expansive vocabulary and an active imagination based on some communicative and sexual literacy. Some participants may also exchange pictures or short movies of themselves or erotic pictures and movies found on the web to accompany the otherwise text-based communication.

TELEVIDEO CYBERSEX

Although televideo cybersex entails considerable text communication as well, there are significant differences between the two. In televideo cybersex, participants can see and respond to one another in real time, by using web cameras and client software. Therefore, in addition to "hot" chat, televideo participants can see each other in live streaming video, including what they look like, what they are (or are not) wearing, where they are, their moment-by-moment expressions, and what they are doing (often to themselves). In this way, televideo cybersex participants can be regarded as simultaneously live erotic performers and consumers, casual voyeurs, and exhibitionists all at once.

As part of sex, stripping and being seen naked is an erotic experience in televideo cybersex. Some participants are excited just because someone wants to see their naked bodies. Dennis Waskul, an expert in cybersex from the field of sociology, suggests that to a certain extent, televideo cybersex can be seen as a loose form of semi-public nudity. Unlike nude beaches, nudist resorts, and Finnish saunas, where nakedness is anti-erotic, the nude body in televideo cybersex is intentionally displayed in an explicitly sexual manner for the purposes of giving and receiving sexual attention and/or arousal. For the majority of participants, televideo cybersex involves a kind of looking-glass eroticism. The viewing pleasures for televideo cybersex participants include both the live moving images of the bodies of other people and the interactive exhibitional imagination. The interactivity of being watched is exciting, fun, and erotic for televideo cybersex users.

Different from the for-profit live webcam performances or services, televideo cybersex refers to those casual and usually anonymous televideo sexual encounters among participants who do it for enjoyment, for sex, and for free, without considering internet service provider (ISP) fees, cost of the digital camera, and cost of the client software. Couples, either heterosexual or homosexual, sometimes connect with each other or other couples in televideo cybersex, but more often, these erotic encounters occur between people who do not know one another and whose interaction is merely for the purposes of casual and immediate sexual gratification.

CYBERSEX AND REAL-LIFE SEX

With the transmission of sexual fantasies via the Internet, for some individuals, cybersex becomes the online equivalent of engaging in sex in real life. Cybersex is such a distinctive kind of sexual activity that it seems to contradict the conventional definition of sex with interaction between physical bodies. The major difference between cybersex and real-life sex is that participants are making virtual love, since it is a cognitive and emotional exchange rather than a physical one.

Different from real-life sex, individuals do not interact directly with the bodies in cybersex. Therefore, the physical barriers which can be found in real-life sex are not present in cybersex. Participants engaging in cybersex create new configurations of sexual experience within the interaction through words and images. By taking away the physical aspect of sex, cybersex is an act of communication with both sides participating to keep the fantasy alive; therefore, it is usually less one-sided than many sexual relationships in real life.

Generally, cybersex can be more spontaneous and adventurous than real-life sex. It is often much easier

to meet someone online in cyberspace than in real world offline. In the real world, it may take days, weeks, or months to cultivate a sexual relationship, while things progress more rapidly online. People often get to the point much faster in their cybersexual encounters, and take more risks than in a traditional face-to-face setting.

Cybersex can be a way for some participants to express themselves sexually in a way in which they may not feel comfortable in a real-life relationship. Anonymity regarding self-disclosure and intimacy is a key factor to consider when studying differences or similarities between cybersex and real-life sex. Typically, persons who engage in cybersex find each other on the Internet and have never met before in real life. Participants feel less inhibited to express their sexual desires due to the anonymity, and they collaborate on a rich sexual fantasy.

Cybersex allows real-life partners who are physically separated to continue to be sexually intimate. In geographically separated relationships, cybersex can have an important function in sustaining the sexual dimension of a relationship in which the partners see each other only infrequently face to face. Although some cybersex participants may develop long-term relationships or even intimate real-life relationships offline, most of the time cybersex experience is anonymous, fleeting, and casual.

APPEALS OF CYBERSEX

The Internet has several characteristics that make it an attractive venue for sexual pursuits. It is widely accessible, affordable, legal, available in the privacy of one's own home, anonymous, and does not put the user at direct risk of contracting a sexually transmitted disease. Other factors include convenience, escape, and social acceptability. In addition, the ability of people using the Internet to approximate, or experiment with, different aspects of sexuality or sexual practices online may also be important. The Internet may allow for sexual experimentation in a forum that seems safer. Unlike print, video, and film pornography, cybersex is interactive, through real-time exchanges of words, pictures, and videos. This enables people to experiment with, and experience, an activity closely, without actually engaging in physical sex, to become psychologically and physically (in an autoerotic sense) familiar in the context of a virtual encounter.

Although cybersex may be regarded as virtual compared to those real face-to-face sexual experiences with interactions between corporeal bodies, it can be meaningful and sometimes even highly valued by some participants, more than a novel masturbatory innovation. Cybersex has become so popular because participants regard it as safe sex, full of fantasy appeal, and sometimes with therapeutic values.

SAFE SEX

The absence of physical bodies promotes the perception of cybersex as a safe alternative to real-life sexual encounters. Cybersex can satisfy sexual desires without the risk of sexually transmitted disease or pregnancy. Cybersex participants can have various kinds of sexual activities by just sitting safely behind their personal computers, and without talking to each other if they do not want to. It can be a physically safe way for young people to experiment with sexual thoughts and emotions. In addition, people with long-term ailments including HIV can engage in cybersex as a way to safely achieve sexual gratification without putting their partners at risk.

Anonymity can provide security and be a protector and a liberator. Participants regard the Internet as a safe technological medium to express or explore different aspects of their sexuality in almost complete anonymity. By taking advantage of the anonymity of cybersex, participants feel safe to share sexual intimacy online. Participants can delete or create a new screen name, if any adventure becomes unpleasant or uncomfortable. Therefore, cybersex becomes a safe outlet for interests and desires which are difficult to be expressed in real life. Anonymity can also help unsociable or stigmatized people because there are no social sanctions, embarrassment, or fear of negative reactions. In addition, the distance afforded by computer-mediated technology adds one more element of safety for the cybersex participants, because cybersex can always stop at the keyboard.

FANTASY APPEAL

Cybersex offers a private, safe, and anonymous way to explore fantasies. The excitement from taking risks without anyone finding out can be part of the appeal of cybersex for some participants, and some cybersex participants may find it erotic just by not knowing who is really on the other end. Anonymity as an important element of cybersexual experiences gives cybersex participants greater sexual freedom. It could be sharing secret sexual fantasies or creating an interactive sex novel. Cybersex provides an opportunity for people to explore and express their most private sexual fantasies. Because of the anonymity, cybersex participants live out their fantasies by doing whatever

they want and taking any identity they wish with no inhibitions.

Another important element of the eroticism of cybersex is that it offers opportunities for participants to playfully toy with alternative vicarious experiences, which they would not or could not try in real life, or in other cases have no intention of trying offline. There are endless possibilities in cybersex for identity change, such as a different gender, race, or even species. Cybersex creates an opportunity for participants to try kinds of sexual fantasies, which are normally hidden or controlled in real life. Compared with real-life sexual encounters, it is relatively easy to find someone to share personal sexual fantasies online.

For some participants, cybersex provides a novel mechanism for them to explore and cultivate sexuality. By exploring cybersex, experimenting and receiving real feedback from partners become quite easy due to anonymity. Some cybersex participants claim that they have learned new sexual techniques and turn-ons, which expand their repertoire of sexual behaviors. By acting out of fantasies, cybersex may help some participants' real sex lives.

THERAPEUTIC VALUES

Some cybersex participants think the enjoyable sexual experiences online also have therapeutic values. For example, some individuals may have a long-standing conflict between internal desires and outside pressures, for fear of punishment or social rejection. For some participants, cybersex can be a way to explore their hidden sexual desires, or free their suppressed sexual desires and needs, which they would never express in real world. It can be a tremendous relief for them. Some ethnographic studies of the online bondage, discipline, sadomasochistic (BDSM) subculture suggest that cybersex may be freeing and helpful in terms of identifying and enacting parts of the person that would otherwise never have surfaced.

As another way to connect to people or to achieve intimacy, cybersex experiences can grow to be a source of comfort for some participants. Relationships can be formed online, such as regular cybersexual encounters, and may become more real as they progress from "hot" chat in the virtual world to direct contact in the real world. For some people, the attention they receive in cybersexual experiences can serve to increase appreciation for and assessment of themselves, and make them feel sexy and attractive.

As for some cybersex participants, it is the escape appeal that entices them go online. They go online to escape the demands and responsibilities in real life for a short time. Besides, without the emotional baggage of performance anxiety in real-life sex, the pressure of having to perform or please a partner seems to disappear in cybersex. Some cybersex participants may have deeper problems in their own primary relationships, and cybersexual experiences may improve their sex life in the real world.

RISKS OF CYBERSEX

Although participants point out a variety of sexual and other personal benefits from online sexual experiences, cybersex also has some potential risks. Online sexual problems include the full range of difficulties that people can have when engaging in cybersex. This would include negative financial, occupational, legal, relationship, and personal repercussions from cybersex. These problems may result from a one-time incident or a pattern of excessive involvement, and the consequences might range from feelings of guilt to loss of a job or a relationship.

EMOTIONAL RISKS

In some cases, engaging in cybersex may pose some emotional risks. Cybersex participants may try something that goes beyond his or her normal comfort range. Some online survey found that cybersex seduces some users into crossing boundaries into sexual thinking, fantasy, longing, and activities in which they may do not want to participate. Some cybersexual encounters can be quite psychologically damaging. Moreover, the issue of sexually related Internet crime, such as cyberstalking, sexual harassment, and molestation may also be imported into cybersex. Some researchers have suggested that the use of cybersex may be counterproductive to normal, healthy sexual development in certain individuals.

Cybersex is usually associated with no necessary commitments. Many people participate in cybersex just for enjoyment, without any intention to take the relationship further. While some individuals enjoy cybersex because of the lack of commitment, others may find this disappointing. For someone who has built an intimate relationship with his/her cybermate, a sudden loss may occur when one individual disappears without warning and the relationship ends abruptly. Someone may feel hurt and angry when he/she goes online find his/her cybermate has vanished, leaving no trail to follow. Such a sudden loss or cyber breakup can be emotionally upsetting. Therefore, for those who are seeking a long-term committed

relationship, they need to be careful when getting too emotionally involved.

CROSSOVER INTO REAL LIFE

The speed and capabilities of the Internet when combined with sexuality have produced far-reaching social impacts beyond the confines of cyberspace. Some cybersex activities may progress to offline sexual encounters with other people. One concern that has become a focal point of social and political action is access by minors to cybersex. Some adolescents may try cybersex as a form of sexual gratification by engaging in various kinds of online relationships, which can be dangerous if they lead to real-life activities, such as underage sex.

Cybersex allows individuals to try out sexual fantasies free of sexual boundaries and limits; however, some extreme fantasies tap into taboo areas. Researchers have studied criminal and deviant use of the Internet and posited that the instant gratification of cybersex provided certain reinforcement for the operationalization of sexual fantasies that would otherwise be extinguished. Some scholars warn of the potential for cybersex to be used as a means of luring children into assignations. Online pedophiles typically masquerade as young people.

Cybersex veers away from healthy sexuality when it interferes with normal responsibilities, causes distress, or becomes out of control. Some cybersex users are reported to get further and further into the bizarre, losing interest in their previous sexual activities and partners, and seeking more and more unusual experiences. Researchers have noted the danger of individuals neglecting their real-world relationships by spending increasing amounts of time engaged in virtual relationships. Some participants may be at great risk for developing a cybersex addiction. A number of self-score questionnaires are available to help cybersex participants judge whether their online activities are causing problems.

Cybersexual experiences can undermine or destroy marriages. Debate continues on whether cybersex is a form of infidelity. Many cybersex participants tend to treat cyberaffairs as nothing more than a harmless form of erotic entertainment since it does not involve physical contact; however, many significant others view cyberaffairs as a definite violation of trust and a serious threat to the marital relationship. In some cases, cybersex can cause such great stress that it can lead to a breakup in real-life relationship. The risks to participants' current relationships in real world need to be considered. Some moral and legal issues may also get involved for those engaging in cybersex.

Another potential risk of cybersex is related to privacy issue. Especially for televideo cybersex, there is the possibility for extreme embarrassment for the participants, since the images can be saved, and even reproduced and distributed. Therefore, with the vulnerability of showing one's face, some televideo cybersex participants choose to keep their face conspicuously absent from images of their body, otherwise cybersexual experiences may carry much emotional weight or pressure in their real life. In addition, the anonymous nature of cybersex permits rather cruel pranks. The intimacy of cybersex may in some cases be rudely shattered by pranksters who solicit cybersex, but with the actual intention to post the logs in public. Many guides for netiquette warn against this.

Another privacy concern is false identity. Role playing and misrepresentation is a typical feature of cybersex. The nature of the Internet as a medium makes it very easy to exaggerate or misrepresent the facts. In cybersex, participants develop various virtual identities, distorting important aspects of themselves and even pretending to be of the opposite sex and of a different age. The virtual identities can go from slight exaggerations or variations to complete fabrications. They are often not easy to detect. Some individuals may feel hurt and ashamed by knowing the true identity of their online partners afterward. When online relationships move from online to offline, in many cases, reality frustrates the fantasy.

CYBERSEX PARTICIPANTS

Researchers propose that not all cybersex users are similar. The majority of participants who engage in cybersex report that their use does not lead to problems in their lives. Nevertheless, for a small but significant minority, cybersex generates serious problems. Studies have been conducted to contribute to an understanding of cybersex participants. Studies led by Al Cooper show that there are three general categories for cybersex users: recreational users, at-risk users, and sexually compulsive users.

Recreational users engage in cybersex exploration more out of curiosity, or for entertainment or experimentation purposes. They are usually satisfied with their online activities, without feeling guilty or ashamed of them. These users engage in cybersex as a casual entertainment and not to an extent that it has any serious negative impact on their life. They tend to maintain reasonable levels of involvement, often become bored or indifferent over time, and decrease or discontinue the activities.

At-risk users have no prior history of sexual compulsivity, yet they may become vulnerable when facing the

accessibility, affordability, and anonymity provided by cybersex, and spend substantial time and energy on cybersex activities. For these participants the power of cybersex interacts with certain underlying personality factors and can lead to patterns and behaviors that, without intervention, may develop into online sexually compulsive behavior. Two subtypes have been described: stress-reactive and the depressive types. The stress-reactive subtype is characterized by the tendency to engage in cybersex during times of high stress. These individuals use cybersex as a temporary escape, distraction, or means of coping with uncomfortable feelings that arise from stressful situations. The depressive subtype seeks relief from depression, dysthymia, and/or chronic difficulties accessing any emotions through their use of cybersex.

The sexually compulsive group is composed of those users who have past or present difficulties with sexual issues in their lives. They, like the others, simply find the Internet to be an effective venue by which to pursue their sexual interests. Unfortunately, their cybersex experiences may escalate the problem, often resulting in clear and significant difficulties in their lives. Sexually compulsive users suffer varying degrees of consequences from their pathological cybersex. In some studies, sexually compulsive users are defined sexual by a compulsivity scale combined with time online, which result in distress and interference with their lives. This group report higher levels of distress around their online sexual pursuits.

In addition, studies have been conducted focusing on specific populations affected by Internet sexuality, including women, gay men, older adults, and clergy. However, as for some other populations, such as lesbians, people with disabilities, and so on, little research is currently available.

CYBERSEX ADDICTION

Cybersex becomes a convenient and relatively safe way to meet powerful psychological forces, such as sex, intimacy and romance; however, the online thrill is potentially addictive. As Internet usage continues to increase, clinicians and therapists report a growing number of clients addicted to cybersex, a form of both Internet addiction and sexual addiction, with the standard problems associated with addictive behaviors. Intoxicating, isolating, integral, inexpensive, imposing, and interactive are aspects of the Internet that appear to be risk factors for cybersex users who have preexisting sexual compulsions or addictions. The anonymity and convenience of cybersex also makes it easy for participants to fall into compulsive use. Beginning with harmless curiosity, cybersex may escalate into compulsive use. Some cybersex participants have frequent erotic encounters online and may border on addiction, finding themselves pressured to engage in frequent cybersexual activities.

Cybersex addiction also features the hallmarks of dependence, including salience, mood modification, tolerance, withdrawal, conflict, and relapse. Salience occurs when cybersex becomes the most important activity in the individual's life and dominates his/her thinking (preoccupations and cognitive distortions), feelings (cravings), and behavior (deterioration of socialized behavior). For instance, even if the individual is not actually online participating in cybersex, he or she may feel preoccupied with going online for cybersex. Some participants may even day-dream about cybersex and look forward to any opportunity to be online for it. Mood modification refers to the subjective experiences that participants report as a consequence of engaging in cybersex and can be seen as a coping strategy, such as the escape appeal of cybersex. Tolerance is the process whereby increasing amounts of cybersex are required to achieve the former mood altering effects. This may mean more hours online for cybersex, a larger number of partners, more bizarre or riskier activities, or going from virtual to actual sexual encounters. Withdrawal symptoms are the unpleasant feeling states and/or physical effects, for instance, the shakes, moodiness, irritability, and so on, which occur when cybersex is discontinued or suddenly reduced. Conflict means cybersex interferes with participants' occupational, social, and/or recreational dimensions of life, including conflicts between the cybersex participants and those around them (interpersonal conflict), conflicts with other activities (job, social life, hobbies, and interests), or conflicts within the individuals themselves (intrapsychic conflict and/or subjective feelings of loss of control), which are concerned with spending too much time engaged in cybersex. Relapse is the tendency for repeated reversions to earlier patterns of cybersex to recur and for even the most extreme patterns typical of the height of excessive cybersex to be quickly restored after years of abstinence or control.

Some scholars have produced a checklist of warning signs for cybersex addiction. These include: routinely spending significant amounts of time in chat rooms and private messaging with the sole purpose of finding cybersex; being preoccupied with using the Internet to find online sexual partners; frequently using anonymous communication to engage in sexual fantasies not typically carried out in real-life; anticipating the next online session with the expectation of finding sexual arousal or gratification; frequently moving from cybersex to phone sex, or even real-life meetings; hiding online interactions from significant others; feeling guilt or shame about online use; accidentally being aroused

by cybersex at first, and now actively seeking it out when online; masturbating online while engaged in erotic chat; and less investment with real-life sexual partners; and preferring cybersex as the primary form of sexual gratification. Based on the *DSM-IV* symptoms of dependence, many screening tools have been designed to help people identify a possible problem with their online sexual behavior, such as the Online Sexual Addiction Questionnaire (OSA-Q) and Internet sex screening test.

There is growing empirical evidence that cybersex addiction exists. Some researchers may not use the term cybersex addiction, although their descriptions of excessive cybersex among the populations they have observed appear to feature the general components of addiction outlined earlier. There is evidence to support the overlap between real-life sex addiction and cybersex addiction. For some participants, cybersex is a continuation of preexisting sexual addiction. Some describe a rapid progression of a previously existing compulsive sexual behavior problem, whereas others have no history of sexual addiction but become rapidly involved in an escalating pattern of cybersex use once they discovered cybersex. For the latter, cybersex is the first expression of an addictive sexual disorder, one that lends itself to rapid progression, similar to the effect of crack cocaine on the previously occasional cocaine user. Some researchers suggest that issues such as isolation and fantasy contribute to at-risk users becoming compulsive.

Studies find that accessing sex on the Internet has the potential to escalate preexisting sex addiction as well as to create new addictive disorders in previously at-risk users. Progression of cybersex addiction is rapid. People who reported a 10-, 20-, or even 30-year history of low-level compulsive sexual behaviors experienced severe life repercussions within a year or two of going online. People addicted to cybersex and related online relationships can experience compulsive usage which means the individual has an intention to stop but inability to do so despite adverse consequences.

In addition, some studies show that both male and female cybersex compulsive users are more likely to use chat rooms, favoring an interactive modality in their sexual pursuits. It may be that the particularly engaging nature of chat rooms may be a slippery slope for certain at-risk individuals in the development of their cybersex compulsivity, or that certain individuals who habituate to less powerful forms of online sexual activities over time gravitate to chat rooms, or that chat rooms are a transitional step from online sexual interactions to meeting and finding face-to-face partners. In any case, the use of chat rooms for sexual pursuits should be a red flag and something to which clinicians should pay particular attention.

Previous studies, usually based on findings from online surveys, document the adverse effects of cybersex addiction on the participant's self-esteem, emotional state, social life, job performance, finances and, in some cases, legal status, when the cybersex involves children or adolescents. Compulsive cybersex use clearly decreases the user's availability to the family. As for the effects of cybersex addiction on partners, they were found to have feelings of hurt, betrayal, rejection, abandonment, devastation, loneliness, shame, isolation, humiliation, jealousy, and anger as well as loss of self-esteem. Thus, it is not surprising that the partner's addiction has been cited as a major factor in subsequent separations and divorce. The reported adverse effects on children include exposure to online pornography, objectification of women and men, or masturbation, involvement in parental conflicts, lack of attention, and breakup of the marriage.

PREVENTION, INTERVENTION, AND TREATMENT

There are already reports of an increase in the number of counselors and therapists seeing people who come in for problems associated with online sexual activities. The extraordinary expansion of computer technology into more and more lives and into all parts of our lives means that cybersex addiction will increase. Cybersex addiction is an extremely potent addiction that must be treated as such. It is usually not just a single family member affected by cybersex, but it can have far lasting effects on partners, children, siblings, and parents. If Internet sex addiction is a viable concept, there are also implications for treatment. At present, treatment programs for sexual addiction include inpatient, outpatient, aftercare support, and self-help groups. There are also family counseling programs, support groups, and educational workshops for addicts and their families to help them understand the facets of belief and family life that are part of the addiction. Some online resources are also available for help, such as Center for Online and Internet Addiction, Internet Behavior Consulting (IBC), and Cybersexual addiction by Sexual Recovery Institute (SRI). However, at present there are very few outlets for the treatment of Internet sex addiction. Clearly there is a need for establishment and evaluation of treatment strategies for online compulsivity/addiction.

Unlike human sexuality, Internet usage is not an innate human need or drive. However, like television and telephone, it has become an essential feature of modern life. Therefore, total abstinence of computer use is probably not the best approach in the long term given the prevalence of computers and Internet use in everyday life. Treatment modalities developed for

treating other addictions, particularly food and sex addictions, are applicable to treating cybersex addiction. Inpatient and outpatient therapy programs can be modified to include cybersex addiction. Inpatient and outpatient treatments vary in intensity, but both include auxiliary treatment such as 12-step groups, peer support, family therapy, medication, and a continuous emphasis on relapse prevention. As for medications for treatment, some cybersex-based paraphilias or hypersexuality can be managed through pharmacological treatment involving anti-psychotic, anti-depressant, and anti-androgenic medications. Depression is a core experience in all addictions; therefore, depression treatment is a central component in the treatment for cybersex addicts. If there is a strong family history of affective disorder and the patient fits diagnosis of the affective disorder and the cybersex addiction, the person may be a good candidate for antidepressant medication. Serotonin reuptake inhibitors (e.g. Paxil or Luvox) are helpful for dealing with depression, anxiety, and compulsive symptoms.

As public and professional awareness of cybersex usage is raised, it becomes increasingly important to understand, assess, and treat this phenomenon. Therapists need to become more informed about the range of sexual activities available on the Internet, the powerful draw of such activities for many cybersex users, the significant adverse consequences that many cybersex users experience, the need to modify self-destructive or illegal behaviors a top priority in therapy, and the importance of considering cybersex addiction a family problem, and involving the spouse or significant other in the treatment process.

Many cybersex participants identify themselves as sex addicts and relate a history of compulsive sexual behaviors antedating their online sexual activities. Therefore, familiarity with current models of treatment for sexual addiction and compulsivity will play a significant role in successful treatment of cybersex compulsives. One option for intervention is the 12-step recovery groups for sexual addiction and/or compulsivity. These groups often assist clients in modifying their compulsive sexual behaviors and interpersonal relationships. In case studies, many respondents reported that what has helped them is attending 12-step sex addiction meetings, daily contact with a 12-step sponsor, individual and couple counseling, and initially, a 90-day abstinence plan. Ideally, the spouse should also attend a 12-step co-sex addict program and as well as individual and couple counseling. Clinicians should be familiar with 12-step groups that are available as an adjunct to individual and group therapies. Training will be needed to help therapists transfer extant knowledge and interventions from work with other sexual acting out problems to the online world, as well as identify and develop methods specifically geared to cybersex issues. In China, Ran Tao's Addiction Medicine Center, inpatient treatment programs are used to jump-start a period of abstinence and a recovery program, especially for adolescent clients. In addition, there are also various online recovery resources, such as online education centers, support groups, and online tests for self-assessment.

Compared with men, female cybersex addicts have additional challenges when they seek help, such as increased shame about the activities, less societal acceptance of women's sexual and cybersexual behaviors, fewer 12-step meetings where women feel comfortable, and lack of knowledge by therapists about cybersex in general and about women's activities in particular.

The results from clinical experience and review of the literature establish that involvement in compulsive cybersex is not a unitary phenomenon. Previous studies find that addictive disorders tend to have their onset in adolescence or young adulthood, but some cases illustrate that cybersex addiction can arise even in later years. Some may have other compulsive behaviors in earlier years, and the powerful lure of cybersex resulted in their switching addictions from others to cybersex. For those with a prior history of compulsive sexual behaviors, the Internet can escalate their addiction and the unmanageability of their lives. Compulsive cybersex is a complex, multifaceted experience that requires several levels of analysis prior to making recommendations for different components of treatment, including assessment, treatment, resources, relapse prevention, and comorbidities.

Some studies find that most of the cybersex abuse patients are married or living with a partner in a committed relationship and primarily college-educated students and professionals, with the youngest age group having most cybersex experience while the oldest group having the least one. Homosexual men were found more likely to have had cybersex compared with heterosexual men, while sexual orientation was found not to have any significant effects on odds ratios for women. When using the number of sex partners during the last year as an indicator of offline sexual activity, having had more than one sex partner increased the possibility of having had cybersex. The results of some online survey showed that a similar proportion of men and women engaged in real-time online sex with another person and significantly more women stated that their online activities had led to real-life sex encounters.

As for psychiatric comorbidities, behavior individuals presenting with substance use disorders and other process addictions, such as pathological gambling or eating disorder, may reveal a concurrent problem with compulsive cybersex. Many compulsive cybersex

patients present comorbid anxiety and affective disorders. In terms of addictive behavior, male cybersex abuse patients were more likely to report chemical dependence, with more males in recovery for alcoholism and more females in active addiction to various drugs (such as benzodiazepines, narcotics, and cocaine). Two-thirds of the women reported an eating disorder, including compulsive overeating and bulimia.

Important future research areas may involve identifying both risk factors and protective factors among those susceptible to Internet sex addiction. The interplay between such factors is likely to be complex, but future research needs to identify the interaction between individual sociodemographic susceptibility risk factors (e.g. gender, age, ethnicity, etc.), psychological risk factors (e.g. personality type, attitude/belief systems, self-esteem), biological risk factors (e.g. genetic predispositions), situational risk factors (accessibility and availability to Internet services, advertising of Internet services, etc.), and structural risk factors (affordability of Internet services, speed of Internet services, etc.).

SUMMARY

Cybersex, as a form of sexual expression accessed through the Internet, is a phenomenon unknown before the mid-1980s. Some characteristics of the Internet explain the appeal of cybersex, such as accessibility (any time, any where), affordability (as inexpensive as a local phone call, and there are a host of ways to get "free" sex), anonymity (people perceive their communications to be anonymous), and escape-friendly.

In both positive and negative ways, cybersex is changing the way people relate. With a powerful fantasy component, many cybersex participants view it as a highly erotic experience. Cybersex provides participants the freedom of sexual expression and experimentation, a way to learn about new sexual techniques, and a means for exploring various aspects of one's sexuality. Cybersex becomes popular also due to its therapeutic values. Although participants can feel less inhibited in cybersex, it is also associated with negative consequences for some users. Cybersex has raised serious legal and social questions, including issues of access, privacy, and morality. Some of the underlying issues are perennial concerns repackaged in novel ways, while other questions are new.

The sexual freedom provided by cybersex may also lead some participants to cybersex addiction. Some investigators have argued that cybersex use ranges along a continuum, extending from adaptive to problematic. The majority of cybersex users who pursue sexual interests online are capable of limiting the time spent in these activities to reasonable levels but many have problems. Empirical evidence indicates that cybersex addiction, with common features similar to other addictive disorders, does appear to exist for some participants. Anonymity, accessibility, and affordability seem to increase the chances for the Internet to become problematic for either those who already have a problem with sexual compulsivity or those who have psychological vulnerabilities rendering them at risk for developing such compulsivity.

For those whose cybersex activities have crossed into compulsivity, adverse consequences for the user and the family can result. This emergent addiction has been reported to ruin individual lives, careers, marriages, partnerships, and families. These individuals are in need of therapeutic intervention for their compulsive use of cybersex, while compulsive cybersex is a complex experience that has yet to be clearly understood. Therefore, as research develops, there will be more evidence-based treatment strategies for cybersex addiction.

SEE ALSO

Sexual Addiction, Internet: Immersive Virtual Worlds, Behavioral Economic Factors in Addictive Processes, The Cell Phone in the Twenty-First Century: A Risk for Addiction or a Necessary Tool?, Overuse of Social Networking

Further Reading

Adamse, M., Motta, S., 1996. Online Friendship, Chat-room Romance and Cybersex: Your Guide to Affairs of the Net. Health Communications, Deerfield Beach, FL.

Carnes, P., Delmonico, D.L., Griffin, E., Moriarity, J.M., 2001. In the Shadows of the Net: Breaking Free of Compulsive Online Sexual Behavior. Hazelden Foundation, Center City, MN.

Cooper, A. (Ed.), 2000. Cybersex: The Dark Side of the Force. Brunner-Routledge, Philadelphia, PA.

Cooper, A., 2002. Sex and the Internet: A Guidebook for Clinicians. Brunner-Routledge, London.

Cooper, A., Delmonico, D.L., Burg, R., 2000. Cybersex users, abusers, and compulsives: new findings and implications. Sexual Addiction and Compulsivity: Journal of Treatment and Prevention 7 (1–2), 5–30.

Daneback, K., Cooper, A., Mansson, S., 2005. An Internet study of cybersex participants. Archives of Sexual Behavior 34 (3), 321–328.

David, L., Delmonico, D., Carnes, P.J., 1999. Virtual sex addiction: when cybersex becomes the drug choice. CyberPsychology and Behavior 2 (5), 457–463.

Delmonico, D., Griffin, E., Moriarity, J.M., 2001. Cybersex Unhooked: A Workbook for Breaking Free of Compulsive Online Sexual Behavior. Gentle Path Press, Wickenburg, Arizona.

Freimuth, M., 2008. Addicted? Recognizing Destructive Behavior before it is too Late. Rowman and Littlefield, Lanham. Distributed by National Book Network.

Griffin-Shelley, E., 2003. The internet and sexuality: a literature review 1983–2002. Sexual and Relationship Therapy 18 (3), 355–370.

Griffiths, M., 2001. Sex on the internet: observations and implications for internet sex addiction. The Journal of Sex Research 38 (4), 333–342.
Jennifer, P.S., 2000. A qualitative study of cybersex participants: Gender differences, recovery issues, and implications for therapists. Sexual addiction and compulsivity: The Journal of Treatment and Prevention 7 (4), 249–278.
Mark, F., Southern, S., 2000. Compulsive cybersex: The new tea room. Sexual Addiction and Compulsivity 7 (1), 127–144.
Robinson, P., Tamosaitis, N., 1993. The Joy of Cybersex: An Underground Guide to Electronic Erotica. Brady, New York, NY.
Schneider, J.P., 2003. The impact of compulsive cybersex behaviors on the family. Sexual and Relationship Therapy 18 (3), 329–354.
Schneider, J., Weiss, R., 2001. Cybersex Exposed: Simple Fantasy or Obsession? Hazelden Publishing and Educational Services, Center City, MN.
Southern, S., 2008. Treatment of compulsive cybersex. Psychiatric Clinics of North America 31, 697–712.
Waskul, D., 2003. Self-Games and Body-Play: Personhood in Online Chat and Cybersex. Peter Lang, New York.
Young, K., 2001a. Tangled in the Web: Understanding Cybersex from Fantasy to Addiction. Authorhouse, Bloomington, Indiana.
Young, K., 2001b. Cybersex: Uncovering the Secret World of Internet Sex. Carlton, London.

Relevant Websites

www.netaddiction.com – Center for Online and Internet Addiction.
www.cybersexualaddiction.com – Cybersexual Addiction by Sexual Recovery Institute (SRI).
www.internetbehavior.com/index.htm – Internet Behavior Consulting (IBC).
www.sa.org – Sexaholics Anonymous.
www.sexhelp.com – Sex-Help.
www.sca-recovery.org – Sexual Compulsive Anonymous.

CHAPTER

82

Video Game Addiction

Daniel L. King, *Paul H. Delfabbro*, *Mark D. Griffiths*[$]

[*]The University of Adelaide, Adelaide, SA, Australia, [$]Nottingham Trent University, Nottingham, UK

OUTLINE

INTRODUCTION

People can gain many benefits from playing video games. These can be educational, social, and/or therapeutic. However, there is evidence that when played to excess, video games can, in some instances, become addictive, especially online video games where the game never pauses or ends, and can be played at any time of the day. For many, the concept of video game addiction seems far-fetched particularly if their concepts and definitions of addiction involve the taking of drugs. Despite the predominance of drug-based definitions of addiction, there is now a growing number of researchers who believe other behaviors can be potentially addictive even though they do not involve the ingestion of a psychoactive drug (e.g. gambling, exercise, sex, Internet use). Such diversity has led to new, all encompassing definitions of what constitutes addictive behavior.

THE MEANING OF VIDEO GAME ADDICTION

The term video game addiction originated in the early 1980s. Video game addiction refers to a persistent and maladaptive pattern of video game playing behavior. Video game addiction is often regarded to be a type of technological addiction, meaning that the technology has similar addictive effects as psychoactive drugs for some users. This subcategory of addictive behaviors also includes Internet addiction. Whereas many types of addictive behavior are recognized as unique clinical disorders, video game addiction has not been widely accepted. There are two main criticisms of the proposed disorder. The first criticism is that there is currently insufficient empirical research to support the claim that video game addiction is a unique psychiatric disorder. The second criticism is that symptoms of video game addiction may be better accounted for by other

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00082-6

psychopathologies, such as depression. Video game addiction is not currently recognized by mental health authorities such as the American Psychiatric Association and the World Health Organization. Thus, video game addiction remains a contentious and purely theoretical disorder.

The limited empirical research suggests that video game addiction involves a psychological dependency on video games. By definition, the term video game refers to a broad range of electronic games played by manipulating images on a video screen. Video games encompass those played on personal computers, dedicated console systems (e.g. Xbox, PlayStation), handheld consoles (e.g. Nintendo DS, Sony PSP), arcade machines, and other portable handheld devices, (e.g. games played on mobile phones, portable mp3 players, etc.). Video game software varies according to genre and format, but most video games emphasize fast action and strategic play. Video games are understood to be psychologically engrossing, and potentially addictive, because they reward skillful or strategic play with a high score or increased status in the game. While players classified as addicted report very high levels of video game involvement, time spent playing video games is itself not a clinically relevant factor in determining addiction. Relative to other addictive behaviors, the addictive potential of video games is considered to be very low.

FEATURES OF VIDEO GAME ADDICTION

There are a multitude of psychological perspectives on addiction, leading it to be defined in many different ways. However, most models of addictive behavior refer to a persistent and uncontrollable urge to consume a substance or engage in an activity that results in significant personal harm and interpersonal conflict for the user. Thus, video game addiction is often said to be present when individuals have completely lost control over their video game playing and their excessive behavior has had a detrimental effect on all aspects of their lives, and compromises their job and/or educational activities, interpersonal relationships, hobbies, general health, and psychological well-being. These two criteria (impaired control and harmful consequences) are regarded as fundamentally important criteria for addiction. An alternative model of addictive behavior has proposed six features or components of video game addiction. Unless sustained continuously for 3–6 months, these criteria may simply indicate a temporary absorption in video games. The criteria include:

1. *Salience.* This occurs when video game playing becomes the most important activity in a person's life, dominating thoughts (preoccupation and cognitive distortions), emotions (cravings), and behavior (deterioration of normal behaviors). An addicted video game player is obsessed with all aspects of video games and, when not playing, will be anticipating or planning the next playing session.
2. *Mood modification.* This refers to changes in a person's mood state that occur as a result of playing video games. Mood change may involve a subjective feeling of euphoria as well as an increase in physiological arousal (increased heart rate, muscle tension, or shaky hands) or, alternatively, a tranquilizing feeling of calm or numbing sensation.
3. *Tolerance.* This refers to the process whereby increasing amounts of video game play are required to achieve the former mood-modifying effects. Tolerance is evident when players gradually increase the amount of time they spend engaged in video game playing. It could be argued that addicted players build up their tolerance to the point that they end a video game session only when they have become mentally or physically exhausted.
4. *Withdrawal.* These are the aversive mood states and/or physical effects that occur when video game play is suddenly discontinued or reduced. Psychological withdrawal symptoms include feelings of frustration, irritability, and flattened affect. Withdrawal motivates the individual to play video games on a regular basis in order to alleviate the unpleasant feeling states associated with abstinence.
5. *Relapse.* This refers to the tendency for the player to revert back to earlier patterns of video game play, and for even the most extreme patterns, typical of the height of excessive video game play, to be restored quickly after periods of abstinence or moderation. Relapse usually indicates that the individual has lost personal control over their behavior.
6. *Harm.* This refers to the negative consequences of excessive video game play. Harm includes personal psychological distress as well as conflicts with other people (family members and friends) and/or other activities (job, school, social life, hobbies, and interests).

John Charlton, an acknowledged expert on technological addiction, suggests that cognitive salience (preoccupation), euphoria (mood modification), and tolerance may not be reliable indicators of video game addiction but merely signify high engagement, or a type of healthy obsession with video games. Therefore, studies may overestimate the prevalence of problem video game play if high engagement is not properly distinguished from video game addiction. Given these issues of reliability, many addiction specialists maintain that impaired control and harmful effects are the most

appropriate criteria for identifying video game addiction.

PHYSICAL SYMPTOMS

Regular video game use may cause a range of adverse health effects for a small percentage of people, including epileptic seizures, motion sickness, headaches, dry eyes, muscle pains, various repetitive strain injuries, and auditory hallucinations. These relatively minor health complaints are often resolved by reduction of or temporary abstinence from video game play. However, video game addicts may experience more serious and ongoing difficulties due to repetitive use, such as carpal tunnel syndrome and migraine headaches. Other physical symptoms of video game addiction include highly irregular sleeping patterns, including insomnia and hyposomnia that impairs the individual's concentration and attention span. Research studies have shown that longer hours of video game play are associated with general ill-health, poor sleep quality and decreased performance at school and work. In some extreme and isolated cases, individuals have died from heart failure following prolonged video game use.

PREVALENCE

The prevalence of video game addiction is influenced by both the general availability of video game machines and the availability of high-speed broadband connectivity. Highly developed urban cities tend to show higher prevalence rates for video game addiction as compared to regional contexts. To date, very few community-based studies of prevalence have been conducted. Prevalence studies have tended to use special populations, such as high school or university students, making it difficult to generalize the results to other populations. Furthermore, most online studies have used self-selected samples that may artificially inflate prevalence rates. Available research has estimated that the prevalence of video game addiction ranges from 0.3 to 2.3%, although prevalence rates in some areas (e.g. South Korea, Japan) have been reported to be as high as 5%. Higher prevalence rates, ranging from 8 to 15%, have been reported in populations of adolescents and college students, particularly those who play online-enabled video games, such as massively multiplayer online role-playing games (MMORPGs). While these prevalence rates are comparable to and higher than other addictive behaviors, very few video game addicts seek out psychological treatment for their video game playing in their lifetime.

COURSE AND ONSET

Video game addiction typically begins in early adolescence. Early indicators or warning signs of video game addiction include spending most or all available hours playing a video game, diminished job and/or educational productivity, choosing to play video games rather than socialize and/or sleep, lying or being secretive about video game use, and avoiding or no longer participating in other activities in order to spend more time playing video games. Although a few individuals may develop symptoms of video game addiction within a short period of time, for most, the course is more insidious. Years of social video game playing may be followed by an abrupt onset, precipitated by greater exposure to video games or by a stressful life event. For instance, children and adolescents can rapidly develop excessive playing behavior following the introduction of a video game machine to their bedroom or family living room. Similarly, young adults who reside on college campuses are at greater risk of developing problematic video game playing habits due to the flexible, unsupervised, and unstructured college environment, unrestricted access to high-speed broadband, and the multiple stressors associated with the transition to living away from home.

In regard to online video game playing, there is generally a progression in the frequency of playing, the planning required to play, and the organization of peers to play with. These video games pose greater challenges and reward the player less frequently in the later stages. Some levels in online video games may require over 4 h of sustained and uninterrupted play with at least five other players. These playing sessions are therefore likely to interfere with real-life commitments and/or activities. The urge to play video games increases during periods of stress, depression, and/or social isolation. Some people sustain an intense and frequent pattern of playing behavior for many years with a single video game, before eventually burning out and then abstaining from playing permanently. For other individuals, problematic video game playing behavior may manifest itself more irregularly, acting as a coping strategy in response to stressful life events.

DEMOGRAPHIC PROFILE

The typical video game player is a male, between 25 and 30 years old, who plays between 6 and 8 h per week, and who has played video games for over 10 years. Video games are also highly popular among children aged 8–15 years, the majority of whom play between 30 and 60 min per day. Other player demographic profiles have recently emerged. In the last

decade, the video game playing market has diversified following the introduction of casual puzzle games, mobile phone games, and online multiuser games that offer more socially driven playing experiences. These video games have attracted a significant female audience as well as players over the age of 35 years. The most diverse video game playing clientele is found in online multiuser video games, some of which attract millions of players. Males are at least twice as likely as females to play video games during adolescence, and males are more likely to maintain their video game playing into adulthood. Males also report playing video games for at least twice as long as females in an average playing session. Video game addicts in most countries tend to be single men in late adolescence or their early adulthood. Female video game addicts tend to be older than male addicts. Female addicts are also more likely to report a concurrent psychological problem, such as depression or anxiety. With regard to other technological addictions, females report greater engagement with non–video game-related Internet activity, including social networking and online shopping, and they are overrepresented in terms of pathological Internet behavior. The reasons for the lack of female participation in video games include the social stigma attached to female video game players, the hostile and overbearing online behavior of some males toward female players (i.e. females are not often socially rewarded for playing), and the masculine imagery and game-play mechanics in video games themselves (i.e. games are often designed by males for other males).

Cultural differences exist with regard to player preferences for video game activities. In Western countries, including the United States, United Kingdom, and Australia, fast-paced, first-person shooting games are a very popular choice among video game addicts. In Asian countries, like South Korea, Taiwan, and Japan, real-time strategy games tend to be more popular. Specific genres of video games, such as role-playing video games, also differ significantly between Western and Eastern contexts in order to appeal to distinct cultural tastes. The prevalence rates of video game addiction are also significantly higher in Far Eastern countries, such as South Korea and Japan.

INDIVIDUAL RISK FACTORS

Some individuals are more susceptible than others to video game addiction. Individuals with poor time management skills are more likely to engage in persistent and uninterrupted video game play. Dispositional factors that have been linked to increased risk of video game addiction include neuroticism, trait anxiety, sensation seeking, compulsiveness, low self-esteem, and depression. It is thought that these factors moderate the effects of self-regulatory mechanisms on an individual's behavior. For instance, anxious individuals hold self-imposed expectations and beliefs that cause apprehension, fear or worry about their environments. These individuals are attracted to video games because the games alleviate anxiety (fear and panic) by providing a safe and predictable retreat from the real world. Parenting style constitutes another risk factor for video game addiction. Authoritative, emotionally distant and unsupportive parents raise children at risk of developing an attachment to video games to satisfy friendship and self-esteem needs. Case study evidence suggests that adolescent video game addicts initially use video games as a maladaptive way of coping with social pressures (bullying or lack of social support), or with unreasonably high expectations to succeed at school. Older video game addicts often report using video games to escape from feelings of dissatisfaction toward work or a close relationship. Video game playing escalates over time as these problems worsen or remain unresolved. Psychophysiological research suggests that excessive computer game playing is maintained through sensitization of the mesolimbic dopaminergic system (the brain pathway that activates feelings of reward), along with a motivational component of reward of specific computer game-related cues. This means that individuals addicted to video games become more sensitized to the cues and rewards of video games over time, making it increasingly difficult to reduce their involvement.

VIDEO GAME STRUCTURAL CHARACTERISTICS

Some video games may have greater addictive potential than others due to their structural characteristics. Structural characteristics refer to the features of the video games that initiate, develop, and maintain player involvement over time. Video game characteristics vary between different video game systems or platforms. For example, the within-game features of arcade video game machines differ significantly from dedicated video game consoles. Early forms of video game play were conceptualized as a nonfinancial form of gambling, due to the structural similarities between video game arcade machines and electronic gambling machines. Both types of machines feature colorful graphics and sound effects, in-game rewards and bonuses for winning moves, digitally displayed scores of correct behavior, and a rapid span of play, negotiable to some extent by the skill of the player. Arcade machines also feature a high-score board, or a socially ranked score that encourages players to compete and outscore their opponents. Studies in the

1980s identified a minority of adolescents who played pay-to-play arcade machines on a regular basis, sometimes absconding from school to play. These players spent disposable income, including lunch and travel money, on the machines, and would often steal money from their parents and/or others in order to finance their video game playing habit. Using a standardized test of pathological gambling (with the term gambling replaced by video game playing), researchers found that some of these adolescent players met the criteria for addiction.

Modern video games are more structurally complex than early arcade machines. Some researchers have proposed five main categories of features to explain the psychological appeal of video games. These include: (1) social features (i.e. the socializing aspects of video games, such as communication features that facilitate social networking), (2) manipulation and control features (i.e. the ways in which a player can interact with and control in-game properties using a physical control scheme and create a sense of mastery and control over the game), (3) narrative and identity features (i.e. the ways in which the player can take on another identity in the game and become a part of an interactive storytelling experience), (4) reward and punishment features (i.e. the ways in which players are reinforced for skillful play (i.e. winning) and punished for losing), and (5) presentation features (i.e. the esthetic qualities of a video game, such as how the game looks and sounds to the player). The reward and punishment features have the strongest association with excessive video game playing behavior. Similar observations have been made in problem gambling research, specifically in regard to electronic gambling machine play. Operant conditioning theory suggests that the frequent and variable reinforcement schedules in video games make them highly psychologically engrossing and resistant to behavioral extinction. The expectation of large rewards and intermittent smaller rewards leads to rapid and persistent playing behavior. Many video game addicts repeatedly perform a small action in a video game in order to obtain a large reward over time. This practice is known among players as grinding, and it can take up a large portion of their video game playing time.

The social aspects of video games can also influence excessive playing behavior. Research has consistently found higher rates of video game addiction in online multiuser video games as compared to single-player video games. A study in the United States found that over 80% of players of first-person shooting games belong to a regular playing group (or clan), and players were primarily motivated by the opportunity to compete against other players. MMORPGs differ from other types of games in that they feature a large interactive game world shared by hundreds, sometimes thousands, of other players. In MMORPGs, players assume the role of a character (usually in a fantasy or science-fiction setting) and take control of many of that character's actions. These actions may involve completing quests or missions (alone or in the company of others) in return for in-game items or other rewards, or competing against other players in organized competitions. MMORPGs are distinguished from other types of video games by the number of players, and by the game's persistent virtual world, which continues to exist and change while the player is away from the game.

In many video games, particularly MMORPGs, players can join social groups (clans or guilds) that enable them to compete against or cooperate with others. The social obligations of these online groups and the unpredictable nature of the video game itself require players to spend long, uninterrupted periods engaged in the video game. Those individuals who fail to play the game at required times or exit the game prematurely may incur an in-game penalty or be ostracized by other members of the group. Therefore, MMORPG addicts often experience significant social pressures to play regularly and limit their absence from the video game. About 9% of the total video game playing population plays MMORPGS, but the prevalence rate of video addiction is reported to be as high as 15% in samples of MMORPG players.

ASSESSMENT

While psychometric scales for assessing video game addiction are currently in use for research and assessment purposes, many have not been clinically validated. Studies of video game addiction in the 1980s used an adapted version of the DSM-III-R criteria for pathological gambling, with the term video game playing substituted for gambling. Individuals who met four or more of the nine criteria were classified as pathological video game players. Some researchers have criticized this approach for its assumption that problem video game playing can be categorized as an impulse disorder. Further, some pathological gambling criteria do not translate easily to video game playing behavior, such as chasing financial losses and increasing gambling bets. A number of self-report measures of video game addiction have been developed since the 1990s. Some of these measures have been based on the updated DSM-IV-TR criteria for pathological gambling, or the ICD-10 criteria for problem gambling, or self-report measures of other addictive behaviors. However, these tests have generally not been cross-validated using a clinical sample of persons classified by trained mental health professionals as addicted to video games. Some researchers have therefore argued that many tests for assessing video game addiction misclassify individuals

as addicted when they are not. This may explain the relatively high prevalence rates of video game addiction reported in some studies.

TREATMENT

Unfortunately, there are currently no evidence-based treatments for video game addiction. There are some specialist addiction treatment clinics (e.g. in The Netherlands, China, Korea, United States) but details of the therapeutic programs have not been published in the academic literature. It would appear that most of the treatment clinics utilize a diverse range of interventions, as do those individual practitioners who have written on the issue (in academic journals and Internet published articles). Cognitive behavioral approaches, in conjunction with family and marital therapies, are often cited as appropriate for treating video game addiction. Therapeutic techniques such as self-monitoring, galvanic skin response biofeedback-assisted relaxation training, in vivo exposure, and response prevention techniques have been used effectively to reduce video game overuse. Since there is very little empirical research into the long-term effectiveness of these programs, it is difficult to establish their credibility. There is currently no proven role for psychotropic medication, although pharmacological treatment has been clinically trialled for the treatment of Internet addiction. A study in the United States in 2008 found that the antidepressant drug escitalopram significantly improved symptoms of impulsive-compulsive Internet usage disorder. Escitalopram has been used previously to treat major depressive disorder, social anxiety disorder, and panic disorder, as well as addictive behaviors including pathological gambling and kleptomania. Some less-conventional treatments for video game addiction have been documented in the popular news media. Most notably, electric shock therapy for video game addiction was being administered to adolescent patients in military hospitals in China, and this practice was discontinued in 2009. Overall, further research is required to establish the therapeutic efficacy of any and all treatment programs directed at excessive gaming.

RESEARCH ISSUES

Research into video game addiction is in its infancy and therefore has a limited knowledge base. A large proportion of video game addiction research involves self-report surveys of players from within university student populations. There is a lack of longitudinal research in this area, which has made it difficult to establish empirically whether excessive video game use represents a stable problem behavior that is maintained on a long-term basis. Furthermore, there is a paucity of in-depth qualitative research on the unique features of video game addiction, including the cognitive and behavioral mechanisms underlying the behavior. Many researchers have criticized the various measures and criteria used to assess video game addiction. The main problems with tests used to identify video game addiction is that they often (1) have no measure of severity to differentiate clinical from subclinical cases, (2) have no consistent temporal response frame, (3) overestimate the prevalence of problems, and (4) take no account of the context of video game use in the person's life. Furthermore, the common practice of using convenience samples in video game addiction research, including first-year university students and users on video game-related Internet sites, may give rise to issues of validity and generalizability of findings. The main disadvantages of self-report methods in video game addiction research include social desirability biases that affect participants' responses to sensitive questions and issues of veracity and reliability of responses due to poor memory recall.

There are some advantages of online self-report methods targeted at video game players including (1) ease of participant recruitment and confirmed presence of the target group under study, (2) collapsed geographical boundaries that may increase numbers of participants in the target group, (3) improved time and cost efficiencies (i.e. allows relatively large-scale samples to be surveyed quickly and efficiently at a fraction of the cost of pen and paper equivalents with no travel needed by either the researchers or the participants), and (4) facilitated data collection and manipulation (e.g. automated data inputting). Online surveys can be particularly useful in gauging opinions from a target group at any particular point in time. The nature of the online medium fosters a relatively high degree of anonymity can be maintained, and video game players may feel more comfortable and thus more likely to disclose personal information when answering sensitive questions on their computer rather than in a face-to-face situation.

CONCLUSION

Based on empirical research, there appears to be little evidence of serious acute adverse effects on health from moderate play. Adverse effects are likely to be relatively minor and temporary. Most will resolve spontaneously as a result of reduction in the frequency of play. Excessive players are the group most at risk of developing more serious problems although further research is required to ascertain the long-term consequences of this activity. There is also a need to determine the

incidence and prevalence of clinically significant problems associated with video game play in the broader population. Epidemiological research would aid in addressing this current gap in knowledge. It is also unclear why so few individuals seek psychological help or treatment for problems arising from excessive video game play. It may be that the detrimental effects of video game addiction are less severe as compared to other addictions, such as gambling, meaning that fewer addicted users believe that some form of assistance is warranted. The predominantly male demographic of video game players may also be less prone to seek treatment, as males are known to be much less likely than females to seek help for a psychological problem. Other factors such as cost, convenience, or a perceived lack of compassion or understanding may also deter video game addicts from seeking psychological help. Worldwide, there are currently very few practitioners who specialize in the treatment of video game addiction. With the growing popularity of video games and the expansion of broadband services, video game addiction is a mental health field that is likely to attract further academic debate and empirical scrutiny.

SEE ALSO

Internet: Immersive Virtual Worlds

List of Abbreviations

DSM	*Diagnostic and Statistical Manual of Mental Disorders*
ICD	international statistical classification of diseases and related health problems
MMORPG	massively multiplayer online role-playing game

Further Reading

Charlton, J.P., 2002. A factor-analytic investigation of computer 'addiction' and engagement. British Journal of Psychology 93, 329–344.

Black, D.W., Belsare, G., Schlosser, S., 1999. Clinical features, psychiatric comorbidity, and health-related quality of life in persons reporting compulsive computer use behavior. Journal of Clinical Psychiatry 60, 839–845.

Fisher, S., 1994. Identifying video game addiction in children and adolescents. Addictive Behaviors 19, 545–553.

Gentile, D.A., Walsh, D.A., 2002. A normative study of family media habits. Applied Developmental Psychology 23, 157–178.

Griffiths, M.D., Davies, M.N.O., Chappell, D., 2004. Demographic factors and playing variables in online computer gaming. CyberPsychology & Behavior 7, 479–487.

Griffiths, M.D., 2008. Video-game and Internet addiction. Adolescent Addiction: Epidemiology, Assessment and Treatment. Academic Press, New York.

King, D.L., Delfabbro, P.H., Griffiths, M.D., 2010. Video game structural characteristics: a new psychological taxonomy. International Journal of Mental Health and Addiction 8, 90–106.

Marshall, S.J., Gorely, T., Biddle, S.J.H., 2006. A descriptive epidemiology of screen-based media use in youth: a review and critique. Journal of Adolescence 29, 333–349.

Shaffer, H.J., Hall, M.N., Vander Bilt, J., 2000. "Computer Addiction": a critical consideration. American Journal of Orthopsychiatry 70, 162–168.

Wood, R.T.A., 2008. Problems with the concept of video game "addiction": some case study examples. International Journal of Mental Health and Addiction 6, 169–178.

Relevant Websites

http://www.netaddiction.com – Internet addiction clinic.
http://www.computeraddiction.com – computer addiction services.
http://www.video-game-addiction.org – video game addiction.
http://www.smithandjones.nl/en/home.html – when it is time for a change.

CHAPTER

83

Exercise Dependence

Jeremy Adams

Eclectic Consulting Pty. Ltd., Altona, VIC, Australia

OUTLINE

EXERCISE DEPENDENCE DEFINED

Exercise dependence represents a behavioral dependence on regular exercising. Exercise dependents typically exercise excessively, and this excess often results in physical and/or psychological and emotional harm. Commonly, diagnosis of exercise dependence requires that a person exercise daily without taking rest days, be likely to have experienced illness or injury as a result of his or her exercising, exercises intensely and for long durations, spends a large proportion of his or her time planning and thinking about exercising, exercises stereotypically (e.g. at the same time or in the same way) and, when unable to exercise, experience irrational thoughts and feelings (e.g. "If I don't exercise today I'll get fat"), as well as withdrawal symptoms (e.g. mood swings, aches and pains, dry mouth).

Despite a lack of evidence regarding the prevalence of exercise dependence, it appears that the condition is relatively rare, with approximately 1% of exercisers experiencing serious consequences from their excessive exercising. Nevertheless, upward of 10% of exercisers exhibit symptoms, presenting a risk for progression to a more serious condition and its potential consequences (see below). The most potentially serious consequence of exercise dependence on a person's health is overtraining, a physiological condition that results from excessive exercising without adequate rest between exercise sessions. Other consequences of exercise dependence include psychological, sociological, and vocational harm, including onset of anxiety and depression (especially in the absence of regular exercise), disruption to personal relationships, and the inability to work effectively.

Along with other nonchemical dependencies, exercise dependence is not presently recognized by the American Psychiatrists' Association's *Diagnostic and Statistical Manual* – fourth edition-text revision (DSM-IV-TR). Despite some controversy (see below), it is widely agreed by exercise-dependence researchers, however, that the condition is legitimate and represents a behavioral dependency rather than, for example, an eating disorder. Consequently, a variety of exercise-dependence diagnostic criteria have been proposed (see Measurement and Diagnosis of Exercise Dependence and Table 83.1).

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00083-8

TABLE 83.1 A Comparison of Contemporary Exercise Dependence Diagnostic Criteria

Author(s)	Diagnostic criteria
de Coverley Veale (1987)*	• Obsession with exercising in a way that is habitual and stereotyped • Noticeable withdrawal symptoms when unable to exercise (e.g. mood swings, irritability, insomnia) • The dependence results in substantial distress or impairment in physical, social, occupational, or other important areas of functioning • The preoccupation with exercise is not better accounted for by another mental disorder (e.g. as a means of losing weight or controlling calorie intake as in an eating disorder)
Ogden, Veale, and Summers (1997)†	Primary exercise dependence • Narrowing of repertoire leading to stereotyped pattern of exercise with a regular schedule once or more daily • Salience with the individual giving increasing priority over other activities to maintaining the pattern of exercise • Increased tolerance to the amount of exercise performed over the years • Withdrawal symptoms related to disorder of mood after cessation of exercise schedule • Relief or avoidance of withdrawal symptoms by further exercise • Subjective awareness of a compulsion to exercise • Rapid reinstatement of the previous pattern of exercise and withdrawal symptoms after a period of abstinence Secondary exercise dependence • As above, but with the associated presence of an eating disorder
Hausenblas and Symons Downs (2002)‡	Clinically significant impairment or distress, as manifested by three or more of the following: **1.** *Tolerance*: defined as either a need for increased amounts of exercise to achieve the desired effect, or diminished effect with continued use of the same amount of exercise **2.** *Withdrawal*: as manifested by either the characteristic withdrawal symptoms for exercise, or the same (or closely related) amount of exercise is engaged in to relieve or avoid withdrawal symptoms **3.** *Intention effects*: exercise is often taken in larger amounts or over a longer period than was intended **4.** *Lack of control*: there is a persistent desire or unsuccessful effort to cut down or control exercise **5.** *Time*: a great deal of time is spent in activities necessary to obtain exercise **6.** *Reduction in other activities*: social, occupational, or recreational activities are given up or reduced because of exercise **7.** *Continuance*: exercise is continued despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by the exercise (e.g. continued running despite injury)
Bamber, Cockerill, Rodgers, and Carroll (2003)§	Primary exercise dependence **1.** Impaired functioning The individual shows evidence of impaired functioning in at least two of the following areas: (a) Psychological – e.g. ruminations or intrusive thoughts about exercise, salience of thoughts about exercise, anxiety, or depression (b) Social and occupational – e.g. salience of exercising above all social activities, inability to work (c) Physical – e.g. exercising causes or aggravates health or injury yet individual continues to exercise even when medically contraindicated (d) Behavioral – e.g. stereotyped and inflexible behavior **2.** Withdrawal The individual shows evidence of one or more of the following: (a) Clinically significant adverse response to a change or interruption of exercise habits. Response may be physical, psychological, social, or behavioral (e.g. severe anxiety or depression, social withdrawal, self-harm) (b) Persistent desire and, or unsuccessful efforts to control or reduce exercise Secondary exercise dependence • As above, but with the associated presence of an eating disorder Associated features The following features are indicative but not definitive: **(i)** Tolerance (i.e. increasing volumes of exercising required) **(ii)** High volumes of exercising and, or exercising at least once daily

(*Continued*)

TABLE 83.1 A Comparison of Contemporary Exercise Dependence Diagnostic Criteria—cont'd

Author(s)	Diagnostic criteria
	(iii) Solitary exercising **(iv)** Deception (e.g. lying about exercise volume, exercising in secret) **(v)** Insight (e.g. denial that exercising is a problem)
Terry, Szabo, and Griffiths (2004)**	• *Salience* — this occurs when the particular activity becomes the most important activity in the person's life and dominates their thinking (preoccupations and cognitive distortions), feelings (cravings), and behavior (deterioration of socialized behavior). For instance, even if the person is not actually engaged in the behavior they will be thinking about the next time they will be • *Mood modification* — this refers to the subjective experiences that people report as a consequence of engaging in the particular activity and can be seen as a coping strategy (i.e. they experience an arousing "buzz" or a "high", or paradoxically tranquilizing feel of "escape" or "numbing") • *Tolerance* — this is the process whereby increasing amounts of the particular activity are required to achieve the former effects. For instance, a gambler may have to gradually increase the size of the bet to experience a euphoric effect that was initially obtained by a much smaller bet • *Withdrawal symptoms* — these are the unpleasant feeling states and/or physical effects which occur when the particular activity is discontinued or suddenly reduced (e.g. the shakes, moodiness, irritability, etc.) • *Conflict* — this refers to the conflicts between the addict and those around them (interpersonal conflict), conflicts with other activities (job, social life, hobbies, and interests) or from within the individual themselves (intrapsychic conflict) which are concerned with the particular activity • *Relapse* — this is the tendency for repeated reversions to earlier patterns of the particular activity to recur and for even the most extreme patterns typical of the height of the addiction to be quickly restored after many years of abstinence or control

* *Developed by author.*
† *Based on earlier work by de Coverley Veale (1987).*
‡ *Based on the DSM-IV diagnostic criteria for substance dependence.*
§ *Developed based on a qualitative examination of 56 female exercisers.*
** *Based on Brown's general components of addiction (Griffiths, 1996).*
Adapted from Adams, J., 2009. Understanding exercise dependence. Journal of Contemporary Psychotherapy, 39(4), 231–240.

To date, there has been very little empirical investigation of the prevalence or incidence of exercise dependence or of potential antecedents, such as age, gender, or activity type, frequency, duration, or intensity. Nevertheless, there is some evidence to suggest that exercise dependence affects around 1% of regular exercisers seriously, while up to 10% might be at risk of developing a dependency on exercise. Similarly, it appears that it takes a relatively long time for exercise dependence to develop to a serious condition; around 2 years of regular, intense, long-duration exercise, although it is not yet understood clearly whether certain (e.g. higher) exercise levels are more likely to exacerbate onset. Because it appears to take some time and quite a lot of effort for the condition to develop, exercise dependence is more common in association with physical activities that encourage regular, long, intense exercise bouts such as running or aerobics classes. Likewise, it is thought that, because these sorts of activities generally involve little social interaction (although aerobics is a group activity, it is usually performed without any interaction with the other people in the class), the lack of an obvious point of reference for a healthy participation level could encourage unhealthy involvement levels. Lastly, although more research is required to determine the effects of gender on risk of exercise dependence, it appears that men are more likely to experience primary exercise dependence (see below) whereas women are more prone to exercise dependence as a secondary aspect of an eating disorder (i.e. secondary exercise dependence, see below). It is possible, however, that these differences simply reflect the prevalence of eating disorders by gender; that is, more women than men experience an eating disorder and, therefore, represent statistically the majority of secondary exercise-dependence cases.

Labeling: Exercise Dependence versus Exercise Addiction

Obligatory, excessive exercising is most usually described as a dependence on exercise rather than as an addiction. This distinction results from the psychological and behavioral consequences of the condition: addiction usually refers to a physiological reliance on a substance or behavior, whereas the term dependence is used to describe the consequences of that reliance, such as the presence of withdrawal symptoms in its absence. As such, exercise dependence is usually described based on the effects of excessive exercising (e.g. the compulsive need to repeat the behavior, withdrawal symptoms, and obsessive thoughts in its absence) rather than the physiological effects. Such a distinction notwithstanding, it is not uncommon for exercise dependence to be referred to by a variety of nomenclature throughout the literature, including

exercise addiction. Other examples have included obligatory exercising, morbid exercising, compulsory exercising, and compulsive exercising.

INVESTIGATION OF EXERCISE DEPENDENCE

Historical

The earliest discussion of exercise dependence occurred in the early 1970s, although, initially, researchers regarded the phenomenon as a "positive addiction", on the assumption that, because exercising had physical and psychological benefits, a dependence on exercise should be beneficial. Subsequently, researchers identified cases of psychological and physical harm resulting from excessive exercising, especially in instances where individuals appeared to be exercising in an obligatory fashion. Early descriptions of this behavior were mostly gleaned from studies of runners and labeled in a variety of ways (see above).

Simultaneously, researchers investigating eating disorders identified cases in which patients diagnosed with an eating disorder were also found to be exercising excessively. These instances were labeled "activity-based anorexia" and a relatively large body of literature emerged describing the link between disordered eating and excessive exercising. Little attempt was made, however, to integrate research into exercise dependence with investigation of excessive exercise in association with eating disorders. This separation resulted in a schism within the literature, with subsequent confusion as to whether persons exhibiting a dependence on exercise were "addicted", or manifesting aspects of an eating disorder. In fact, within the eating disorder research, there has been a large amount of skepticism as to whether exercise dependence actually exists independently of an eating disorder, an argument that, to a lesser extent, still continues (see below).

One area in which the majority of these early studies corresponded was their lack of consensus over the condition itself. Not only was there a lack of agreement of the terminology (see above), there was little agreement on the possible antecedents to the condition, let alone its measurement or diagnosis.

Contemporary

In the late 1980s, researchers attempted to align competing ideas about exercise dependence versus manifestations of an eating disorder by proposing the notions of primary and secondary exercise dependence. Briefly, according to researchers, secondary exercise dependence represents a dependence on exercise in combination with disordered eating, whereas primary exercise dependence describes exercise dependence in conjunction with normal or healthy eating (i.e. in the absence of an eating disorder). Both of these terms will be explored in greater detail below. Subsequently, a number of researchers expanded on these terms by investigating both conditions and providing some (albeit controversial) empirical support for the existence of both. As discussed below, however, there is still some contention over the existence of primary exercise dependence as a pathological disorder, as well as a distinct lack of research regarding the risk factors for exercise dependence, its etiology, and its treatment.

NEUROBIOLOGICAL EXPLANATIONS

There has been limited investigation of the potential neurobiological antecedents and correlates of exercise dependence. Nevertheless, several researchers have proposed a variety of mechanisms. Early explanations implicated the endogenous opioid systems, suggesting that exercisers can achieve a "runners' high" through exercise-induced elevations of β-endorphins in the bloodstream. Unfortunately, there appears to be little empirical evidence for this hypothesis. Not only is it unlikely that sufficient levels of β-endorphins would pass through the blood–brain barrier to have a noticeable effect, there have been no reports in investigations of human exercisers that evidence the opiate-induced euphoria attributed to the so-called runner's high. As well, investigations of endogenous opiate levels in sacrificed mice induced to run excessively have shown no evidence of excessive activation of the endogenous opiate systems.

Another proposed explanation for exercise dependence, labeled the "sympathetic arousal" theory, suggests that modifications to the expression of the catecholamines epinephrine and norepinephrine might be sufficient to modify exercise behavior. According to this theory, regular exercising results in greater metabolic efficiency, resulting in a reduction in the basal expression of catecholamines. Because the catecholamines are implicated in general arousal, increases in fitness would result in lower overall catecholamine levels and, therefore, lower arousal. Thus, in order for an individual to maintain his or her arousal levels within a normal range, he or she would need to exercise at increasingly higher intensities and for longer durations and, correspondingly, would experience a substantial drop in arousal levels when deprived of regular exercise. Although this theory is based on sound conceptual reasoning, to date, there have been no empirical investigations either in humans or animals that have provided any evidence to support the notion that tolerance to

catecholamine expression could explain exercise-dependent behavior.

Of greater promise in explaining exercise dependence from a neurobiological perspective is exercise-based modifications to the mesolimbic dopamine system (MDS). The MDS (sometimes referred to as the brain's "pleasure center") is a midbrain pathway that includes the nucleus accumbens and the ventral tegmental area. This system appeared relatively early on in mammalian evolution and exists primarily to reward behaviors related to survival through the expression of dopamine (e.g. feeding, sexual behavior, maternal and paternal behaviors). Among other processes, dopaminergic expression results in subjective feelings of pleasure and contentment, and the experience of these sensations often results both in activation of an emotional response and long-term memory storage of the behavior in question (in order to ensure its repetition). Unfortunately, for humans, the MDS can be activated relatively easily by a range of behaviors, making it possible for nonessential behaviors (e.g. gambling, sex, overeating, exercise) to be experienced as essential. In a similar vein, chemicals that affect dopaminergic cells can effectively 'fool' the midbrain into encouraging their repeated consumption. For instance, cocaine can act as selective dopamine reuptake inhibitor, increasing the amounts of available dopamine and, therefore, reinforcing ongoing cocaine consumption.

Because the brain is able to self-modify in response to environmental and biological changes, increases in available dopamine result in compensation in the MDS to reduce the availability of post-synaptic dopamine receptors. This modification allows the brain to achieve a balanced level of dopamine when levels are raised artificially (e.g. through ingestion of chemical substances, or by engaging in a behavior that stimulates MDS dopamine expression). Problems occur when the substance or behavior that induced the change is withdrawn abruptly. In these instances, MDS modifications to reduce the number of receptors, combined with a reduction in dopamine expression results in substantially lowered available dopamine and can lead to a number of physiological and psychological effects. Most notably for humans, a rapid reduction in available dopamine levels can result in strong cravings for the substance or behavior responsible for elevating dopamine levels artificially.

A second problem with MDS modifications to stabilize dopamine expression following use of a substance or engagement in a behavior, is that the user becomes progressively unable to experience pleasure in the absence of that substance or behavior. In other words, without regular access to the stimulating chemical or activity, other previously pleasurable activities (such as eating, social interaction, sexual behavior, or esthetic stimulation) become decreasingly desirable resulting, eventually, in anhedonia: an inability to experience pleasure except when exposed to the chemical or behavior responsible for the MDS modifications.

There is a relatively large body of evidence that supports the dopaminergic modification theory as a factor in both chemical and behavioral dependencies, as well as in exercise dependence specifically. For example, in comparison to controls, when volunteers were asked to play a video game for monetary reward they showed elevated dopamine expression in the MDS, similar to that seen following amphetamine consumption. Similarly, human exercisers who have ceased exercising voluntarily for a period of time have reported withdrawal symptoms similar to those seen in withdrawal from substances that affect the MDS directly (e.g. cocaine, amphetamines, methamphetamines). More convincing evidence comes from animal models. Investigation of the MDS in mice and rats sacrificed after being induced to exercise excessively (through either running or swimming) has shown substantial reductions in the number of available post-synaptic dopamine receptors, suggesting that exercise can, in fact, modify the MDS over time. Combined with human-based studies which have indicated increased dopamine production post-exercise (especially at high exercise intensities of longer duration), and a number of studies that have described acute elevations in mood post-exercise, it appears that the dopaminergic expression hypothesis provides a relatively convincing explanation for exercise dependence. Nevertheless, until these modifications can be demonstrated directly in human exercisers, this theory will remain speculative.

In addition to the possible activation of dopaminergic systems in exercise dependence, investigation of compulsive voluntary activity in mice when deliberately starved, as well as evidence of a reduction in appetite in mice forced to exercise, has resulted in another possible explanation. According to investigators, this so-called "activity-based anorexia" might represent a dormant survival mechanism that evolved in mammals prior to *Homo sapiens*. Put simply, to protect against starvation during famine, early mammals might have evolved a mechanism that stimulated energy expenditure when calorific intake dropped dramatically. Although counter-intuitive, this behavior might have allowed for migration to an area with better food supplies. Similarly, an automatic reduction in calorific expenditure in conjunction with high-intensity exercise might have allowed for a migrating organism to continue to function without the need to stop and scavenge for food, potentially increasing the chances of finding a food-rich area. In humans, there is certainly a documented relationship between excessive exercising and

disordered (i.e. drastically reduced) eating (see below). Nevertheless, without further convincing neurobiological and behavioral evidence for this theory, it will be difficult to validate.

PSYCHOLOGICAL EXPLANATIONS

Neurobiological explanations can account for a proportion of the factors in explaining exercise dependence. Nevertheless, a variety of psychological variables are also likely to influence the onset, development, and maintenance of exercise dependence. In particular, the positive connotations associated with exercising could be an influencing factor in excessive exercising. Although it is indisputable that regular exercise is of substantial benefit, both physically and psychologically, the mistaken assumption, in a minority of exercisers, that increasing amounts of exercise will result in greater benefits could lead these persons to continue to increase their exercising to an unhealthy level (at which point the physiological processes associated with dependency will have had time to develop). Simultaneously, the social reinforcement associated with exercising (i.e. it is seen widely as a highly positive behavior and rewarded accordingly) could also encourage some exercisers to continue to exercise at unhealthily excessive levels. Speculation notwithstanding, there has yet to be convincing empirical evidence implicating social desirability and reinforcement factors in the onset of exercise dependence.

Another possible psychological factor in the onset and reinforcement of exercise dependence, that has received some empirical investigation and support, is the reinforcing effects of negative withdrawal when exercise is unavailable. A variety of investigators have confirmed the existence of withdrawal symptoms (including mood swings, irrational cognitions, "flu-like" sensations, and anxiety) following exercise cessation in those who exercise excessively. Unfortunately, as with much of the research in this field, there has been little consensus between researchers as to how much exercise participation defines a classification of exercise dependence. Two other problems with this body of research are evident. First, there has been no investigation of the longevity of these withdrawal symptoms and whether they are ameliorated by a reuptake of exercise. Second, because it is difficult to ask those diagnosed with exercise dependence to quit exercising altogether, most of the research on exercise withdrawal has relied on retrospective self-reports from exercisers deprived of their regular exercising by injury. Given that injury can result in psychological trauma the assumption that exercise deprivation and not the injury that explains the supposed withdrawal is premature. Consequently, while it is generally agreed that withdrawal can occur in regular exercisers deprived of their ability to exercise, it is difficult to determine whether these findings relate to exercise dependence *per se*, or simply represent the effects of a forced decline in exercise patterns among committed exercisers.

PRIMARY VERSUS SECONDARY EXERCISE DEPENDENCE

As briefly described above, exercise dependence is best conceptualized as one of two separate disorders: primary and secondary. On the one hand, primary exercise dependence represents a condition in which an individual demonstrates a dependency on exercise (based on diagnostic criteria – see below) in the absence of any eating-related pathology. In other words, in cases of primary exercise dependence, eating is likely to be normal despite excessive, compulsive exercising. On the other hand, secondary exercise dependence represents obsessive exercising in conjunction with disordered eating (such as anorexia nervosa or bulimia nervosa) and might represent an aspect of the eating disorder rather a separate condition. It is not uncommon for exercise to be used as a means of purging or calorific restriction (increasing calorific expenditure to reduce overall calorific balance) in persons suffering from an eating disorder, a fact that has led some researchers to question the notion of exercise dependence in these cases. In other words, because eating disorders often include an element of exercising as a part of the disordered behavior, it is potentially difficult to argue that a dependence on exercise is evident. Such a distinction will require further research. What is not in contention is that eating disorders, with or without an exercising component, invariably result in psychological and physiological harm, and are often associated with comorbid psychopathology, such as depression and anxiety.

Because it has yet to be determined whether it is likely to coexist with psychopathologies, there has been some contention as to whether primary exercise dependence can be classified as a pathological condition. Some researchers have argued that diagnosis of dependence requires that the dependent behavior be associated with some sort of psychological harm, such as psychopathology. For instance, investigation of exercise dependence in female exercisers has resulted in identification of excessive exercising in association with an eating disorder (i.e. secondary exercise dependence) alongside evidence of substantial psychopathology in this group (e.g. DSM-IV TR Axis 1 disorders). These investigators were unable, however, to identify any evidence of psychopathology among those who exercised excessively but who did not show signs of an eating disorder (i.e. primary exercise dependence). Nevertheless, the

use of an exclusively female sample to assess exercise dependence is likely to result in higher rates of eating disorders than one including male participants (reflecting the substantially higher rates of eating disorders among women). Thus, given the finding that, in women, excessive exercising is usually accompanied by disordered eating, it is unlikely that primary exercise dependence would be detected readily in an all-female sample. Subsequently, other researchers have suggested that primary exercise dependence might be more common in male exercisers than in females. Although this notion has received anecdotal support, it still requires empirical demonstration. Future researchers also need to determine whether psychopathology can be detected in persons who exercise excessively and who show signs of dependence.

Psychopathology is not, however, the only form of harm that can result from a dependency. Whether or not psychopathology exists in primary exercise dependence, as is discussed below, there are a number of other potential harmful consequences. Consequently, some researchers have suggested that primary exercise dependence can be classified as a pathological condition in its own right.

POTENTIAL NEGATIVE CONSEQUENCES OF EXERCISE DEPENDENCE

Of particular importance in understanding exercise dependence is the notion of harm as a consequence of the excessive behavior. As with other dependencies, an important aspect in the diagnosis of exercise dependence is that the behavior must have resulted in some form of dysfunction, either psychosocially or physically. Some of these potential outcomes are explored below.

Psychosocial

As with other forms of dependence, excessive exercising is likely to result in reduced psychosocial functioning as the behavior begins to dominate a person's time and other resources. The increasing amounts of time required to sustain an exercise habit is usually associated with a withdrawal of time and energy from other aspects of a person's life, such as his or her commitment to social and romantic relationships, and work and recreational activities. Unlike some other dependencies in which it is possible to obtain a hedonic effect rapidly (e.g. amphetamine use), a large amount of exercise is required to sustain an exercise habit. Although there are no direct investigations of the amount of exercise performed typically in exercise dependence, anecdotal reports indicate that many exercise dependents exercise in excess of 3 h day^{-1}, and spend a large amount of their remaining time planning exercise or engaging in other exercise-related activities.

The main consequence of the large amount of time and energy required to sustain exercise dependence is a large reduction in interpersonal interactions, unless those interactions are exercise related (e.g. relations at a running club or gym). As such, few activities that might have been considered pleasurable pre-dependence are likely to continue. Combined with the likely anhedonia (inability to experience pleasure) that accompanies most dependencies, there is little motivation for those with exercise dependence to engage in activities unrelated to their exercising. Unfortunately, as with many other areas of exercise dependence, there has, to date, been little empirical investigation of its psychosocial consequences.

Physiological

The main physical risk factor from exercise dependence is overtraining, a condition in which an individual exercises to excess, resulting in systemic physiological damage. Physical fitness is, essentially, a physical adaptation to exercise-induced stresses. When exposed to the damage resulting from exercising, over a 24–48-h period without repeated stress (i.e. rest) the body adapts by "overcompensating" in its repairs, slightly increasing the ability of the stressed systems to cope with future trauma (i.e. resulting in stronger muscles, connective tissue and bone, increased blood supply). Over time, these adaptations can result in a substantial increase in fitness levels. If an individual continues to place stress on his or her body, however, by exercising without adequate rest between exercise sessions, he or she will not be able to adapt, and will eventually succumb to exercise-induced damage (i.e. overtraining). Overtraining can result in harm to a variety of physiological systems, including the musculoskeletal, endocrine, cardiovascular, reproductive, and immune systems. As such, overtraining can result in injury, illness, and permanent damage to an individual's ability to reproduce, or to regulate important physiological systems (such as the hypothalamic-pituitary axis, HPA). There have been anecdotal reports of chronic fatigue syndrome (CFS) in association with athletic overtraining and in cases of exercise dependence induced overtraining. Unfortunately for exercise dependents, whereas overtraining in athletes is often detected in time, and modifications made to their training and competitive schedules in order to reduce or reverse the effects, excessive exercisers seldom recognize the signs of overtraining and continue to exercise without reductions in or

modifications to their training, potentially resulting in serious, long-term physiological damage.

MEASUREMENT AND DIAGNOSIS OF EXERCISE DEPENDENCE

A wide variety of instruments purported to measure exercise dependence have been presented in the literature over the last 40 years. As discussed above, researchers investigating exercise dependence have often failed to agree on a definition of the condition, or a set of diagnostic criteria and, consequently, the majority of measurement instruments reflect this lack of consensus. Similarly, until relatively recently, few measures of exercise dependence have been shown to possess adequate psychometric robustness. Thus, it is hard to determine what, exactly, the majority of earlier exercise-dependence measures were actually measuring. Perhaps the most relevant criticism of the majority of exercise-dependence measures is the lack of a set of population normal scores to which individual responses can be compared. Without this information it is hard to determine a "cut off" criteria to determine whether a person is at risk of exercise dependence or not. As well, the majority of measures to date have been developed for a specific exercising population (e.g. runners), or have not been tested across a variety of exercise types, or exerciser experience or proficiency levels. Without this sort of information it is not possible to use these instruments in a way that allows researchers to compare scores across different types of exercise or exercisers. More recently, several instruments have been developed (most notably the Exercise Dependence Scale Revised: EDS-R), which allow for comparison of an individual's scores to normal tables, allowing researchers to estimate exercise-dependence risk more effectively. Work is still required, however, to validate this measure across a variety of exercising types and proficiencies.

In addition to its measurement, since its conceptualization a variety of criteria have been proposed for the diagnosis of exercise dependence. The bulk of these criteria is summarized in Table 83.1. Essentially, there are two sets of criteria: those that are based on the DSM-IV TR criteria for substance abuse, and those that have been developed independently of the DSM. Nevertheless, despite conceptual variations, when comparing the various diagnostic conceptualizations, they appear similar, most likely the result of a stable agreement between researchers regarding the processes, signs, and symptoms of the majority of dependencies (behavioral and chemical). Of note is the fact that several of the diagnostic criteria presented in Table 83.1 are based on theorization about what should be observed in cases of exercise dependence, whereas the remainder are based on actual observations of individuals who have shown evidence of exercise-dependent behavior.

In essence, exercise dependence is diagnosed when the exerciser exercises often, to excess (i.e. on more days than not, at high intensities and for long durations) and stereotypically (i.e. little variation in training regimen), and demonstrates substantial reluctance to any suggested reduction of the level of exercise undertaken, or to missing regular exercise sessions. In addition, exercise dependents find it difficult to control their urges to exercise, and usually experience physical and, or psychological withdrawal (including unwelcome, intrusive cognitions) when unable to exercise. Equally, their dependence reduces their involvement and ability to function in other important life areas, such as emotional, social, and occupational activities. Lastly, their compulsion to exercise is not better explained by another disorder. In the case of secondary exercise dependence (see above), diagnosis is made when the conditions for primary exercise dependence coexist with an eating disorder.

It is up to the individual practitioner or researcher to decide which diagnostic criteria to ascribe to when attempting to identify exercise dependence, however, the more recent suggestions, such as those of Bamber and her colleagues, or Hausenblas and Symons Downs (see Table 83.1), are based on more rigorous investigation of the condition, and represent the end product of an epistemological process, rather than theorization about how exercise dependence ought to manifest.

TREATMENT OF EXERCISE DEPENDENCE

To date, there has been no empirical investigation of the most effective methods for treating exercise dependence. Nevertheless, given the similar likely etiology to other behavioral and chemical dependencies, it is not unreasonable to expect that treatments that have been effective in their treatment would also be of use in exercise dependence. For example, cognitive behavioral therapy (CBT) has been reported to be highly effective in the treatment of a variety of behavioral and chemical dependencies. Likewise, motivational interviewing, a technique developed to treat chemical dependencies, is also likely to be effective in assisting those with exercise dependence, especially for helping to stimulate behavioral change in those who are reluctant to do so.

Practitioners wishing to treat exercise dependence should make themselves familiar with the various

diagnostic criteria, particularly those based on empirical investigation (see Table 83.1), in order to be able to identify the disorder correctly. Because of the possibility of comorbid disorders in conjunction with exercise dependence (e.g. anxiety, depression, eating disorders), practitioners should also ensure that they possess sufficient training in order to be able to treat these problems, or make sure that they are able to refer to someone with appropriate qualifications and experience.

SUMMARY

Exercise dependence is a disorder in which persons exercise excessively and obsessively to a point of dependence and, often injury or illness. Other characteristics of the disorder include withdrawal symptoms when exercise is unavailable, unwelcome, or irrational cognitions concerning exercise (or a lack thereof), a reduction in social and occupational interaction, and reduced functioning in other areas (e.g. interpersonal relationships). Although, to date, there have been no conclusive studies regarding the neurobiological determinates of exercise dependence, it is likely that it is brought about, over time, by exercise-induced modifications to the MDS in the midbrain, activated by sustained, high-intensity exercise.

Despite wide ranging investigation of exercise-dependence spanning some 40 years, until recently there has been little consensus between researchers over its etiology, diagnosis, treatment, or risk factors. There remains contention between researchers as to whether a primary dependence on exercise can, in fact, exist, despite numerous anecdotal reports of such a condition resulting in harmful consequences. Although there have been useful attempts at describing exercise dependence in a variety of exercising populations, including the development of relatively comprehensive and well-reasoned diagnostic criteria, more rigorous and well-planned research is required to determine whether exercise dependence does, in fact, occur regularly in exercisers, what the likely predisposing factors might be, and identification of the best methods for its measurement, diagnosis, and treatment.

SEE ALSO

Gambling, Internet Addiction: Cybersex, Video Game Addiction, Work Addiction, Sexual Addiction

List of Abbreviations

CBT	cognitive behavioral therapy
CFS	chronic fatigue syndrome
DSM-IV TR	*Diagnostic and Statistical Manual* – fourth edition – text revision
EDS-R	Exercise Dependence Scale Revised
HPA	hypothalamic-pituitary axis
MDS	mesolimbic dopamine system

Glossary

Dopamine a catecholamine neurotransmitter and neurohormone found largely in the mesolimbic dopamine system comprising portions of the midbrain (including the nucleus accumbens and ventral tegmental area).

Exercise any structured physical activity that results in elevated physical load, including raised heart rate, respiration, and energy consumption.

Exercise dependence a behavioral dependency on exercise categorized by stereotypical exercising, a reluctance to take rest days, exercising despite pain or injury, and a compulsive desire to exercise once or more daily.

Primary exercise dependence exercise dependence independent of any majorly disordered eating.

Secondary exercise dependence exercise dependence concurrent with substantially disordered eating.

Further Reading

Adams, J., 2009. Understanding exercise dependence. Journal of Contemporary Psychotherapy 39 (4), 231–240.

Bamber, D.J., Cockerill, I.M., Rodgers, S., Carroll, D., 2003. Diagnostic criteria for exercise dependence in women. British Journal of Sports Medicine 37, 393–400.

de Coverley Veale, D.M., 1987. Exercise dependence. British Journal of Addiction 82 (7), 735–740.

Hausenblas, H.A., Symons Downs, D., 2002a. How much is too much? The development and validation of the exercise dependence scale. Psychology & Health 17 (4), 387–404.

Hausenblas, H.A., Symons Downs, D., 2002b. Exercise dependence: a systematic review. Psychology of Sport and Exercise 3 (2), 89–123.

Miller, W.R., Rollnick, S., 2002. Motivational Interviewing: Preparing People for Change, second ed. Guildford, New York.

CHAPTER

84

Work Addiction

Cecilie Schou Andreassen[*,$]

[*]The Bergen Clinics Foundation, Norway [$]Department of Psychosocial Science, University of Bergen, Norway

OUTLINE

INTRODUCTION

Work addiction [The terms "work addiction," "workaholic," and "excessive work" have sometimes been used in the existing literature to describe the same phenomenon. In this article, I will stick with the term "workaholism," which is consistent with the academic literature on the field of study.], often referred to as workaholism, is defined in many different ways, but most definitions include notions of overindulgence with work, long working hours, working more than what is demanded/expected, prioritizing work over most other activities, enjoyment of work, work compulsiveness, and perfectionism, rigidity, high motivation, resourcefulness, impatience, and self-absorption in work. Thus, the term has both negative and positive connotations. A definition that aims to cover all of the above-mentioned attributes should comprise effective, cognitive, and behavioral elements, as well as positive and less positive aspects.

Since some authors view workaholism in positive terms and others in negative terms, it has been hard to identify one homogenous group of workaholics. Unfortunately, existing typologies rely to a large extent on purely atheoretical rationales that have not been investigated empirically. Scott et al. postulated three types of workaholics: compulsive-dependent, perfectionist, and achievement-oriented. These three types are viewed as stable patterns linked to different antecedents and outcomes, but workaholics can also display a mixed pattern. Robinson proposed four different types of workaholics: the "all or nothing" or "bulimic," who does things perfectly or not at all; the "relentless" who has problems stopping work, is compulsive, and is motivated to work fast and meet deadlines; the "savoring" who obsesses over details to the point of paralysis; and the "attention-deficient" who tends to start many projects but gets easily bored and needs to be stimulated at all times. One central, empirically supported distinction is drawn between workaholics who enjoy their work and workaholics who do not. Spence and Robbins identified these two following types of workaholics: "enthusiastic," who are characterized by high levels of work involvement, driven by an internal pressure to work, and find great fulfillment and pleasure in work; and "non-enthusiastic," who are also highly involved in work-related activities and are driven to work, but who seem not to derive enjoyment from doing so.

ASSESSMENT

In terms of assessment, Spence and Robbins define the concept by identifying three independent dimensions or components that they argue characterize workaholism: high involvement in work (Work Involvement), feeling

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00084-X

driven or pressured to work by an internal motivation (Drive), and lack of enjoyment from doing so (Enjoyment of Work). These three dimensions comprise the Workaholism Battery (WorkBAT), which is clearly the one workaholism measure that has received most empirical attention so far. Evidence for the reliability and validity of WorkBAT has emerged from a number of different sources, based on heterogeneous and cross-cultural samples. The WorkBAT consists of 25 items answered on a 5-point Likert scale ranging from strongly disagree to strongly agree (e.g. "I feel guilty when I take time off work"). The item pool is drawn from a review of existing literature as well as its creators' own hypothesis. To qualify as a workaholic of any kind, there should be high scores on Work Involvement and Drive, whereas scores on the Enjoyment of Work subscale distinguish the two types. Some researchers have followed this dichotomized classification of data when reporting their results, while a few have preferred analyses using the full variance of data. Recently, the original three-factor structure and psychometric properties of WorkBAT have been questioned. Two (Enjoyment of Work and Drive) of the three proposed subscales have consistently met the reliability and validity requirements expected for scales of this type. However, the Work Involvement subscale has failed in several studies to display appropriate psychometric properties, with low internal consistencies and insufficient factor loadings. To address this problem, McMillan et al. have developed and offered evidence for a revised 14-item two-factor (Drive-R, Enjoyment-R) version of the WorkBAT, named WorkBAT-R. A recent study investigating a Chinese version of the WorkBAT in Taiwan actually found support for a five factor solution comprising the following factors: Work Enjoyment, Work Involvement-Enjoyment, Drive-Work Involvement, Drive, and Work Involvement.

However, the first workaholism scale that was developed was the Work Addiction Risk Test (WART). The WART consists of 25 items answered on a 4-point Likert scale ranging from 1 (never true) to 4 (always true) (e.g. "I find myself continuing to work after my co-workers have called it quits"). The construction of the item pool is based on recognized features reported by clinicians working with "workaholics" and their families. High scores on WART (57 and above) indicate workaholism. To a large extent, the item pool covers Type A traits (e.g. impatience, doing multiple tasks simultaneously) and anxiety. When factor analyzing WART, Flowers and Robinson found five separate but correlated factors, which they named Compulsive Tendencies, Control, Impaired Communication/Self-Absorption, Inability to Delegate, and Self-Worth. Several studies, primarily using narrow U.S. samples, have been conducted using WART. Evidence for fairly good psychometric properties has emerged, but the use of a shortened 15-item version in future research has been recommended. Recently, a Dutch version of WART provided support for the original 25-item version and the five factors, but it was concluded that eight items from the Compulsive Tendency (CT) subscale were sufficient to capture workaholism. In a similar vein, another Dutch study reported a strong correlation (0.59) between the Compulsive Tendencies subscale (renamed "Working Excessively") and the Drive subscale from WorkBAT described above. The overall knowledge about the psychometric properties of the WART still remains scarce. The study from Taiwan reported that above Work Enjoyment and Work Involvement were unrelated to the WART sum score, whereas the three other subscales of the WorkBat found in that study all showed positive relationships with the WART sum score.

Another workaholism scale, the Schedule for Nonadaptive Personality Workaholism Scale (SNAP-Work) was developed by Clark. It comprises an 18-item, forced-choice questionnaire. High scores on SNAP-Work are assumed to reflect workaholism. The item pool largely comprises obsessive–compulsive traits (e.g. perfectionism, rigidity, high self-imposed demands). Although the limited amount of studies conducted using this scale has demonstrated good psychometric properties, it has not yet received widespread empirical attention. Significantly, McMillan et al. found a strong correlation (0.61) between SNAP and their revised WorkBAT-Drive subscale.

Recently a new workaholism questionnaire was developed. This is named as the Dutch Workaholism Scale (DUWAS). The DUWAS has two subscales: (1) Working Excessively (WE), which actually consists of five items of the Compulsive Tendencies subscale of the WART and (2) Working Compulsively, which consists of five items from the Drive scale of the WorkBAT. All items are scored along a 4-point rating scale, ranging from 1 (totally disagree) to 4 (totally agree). The DUWAS has shown good psychometric properties in a Dutch and a Japanese sample as well as in a Spanish sample. Interestingly, the DUWAS comprises no items reflecting work enjoyment and the scale constructers regard work enjoyment not to be a constituting element of workaholism.

PREVALENCE

In the wake of globalization, new technology, blurred boundaries between work and other life domains, some authors have suggested that we are witnessing an increase in workaholism. Concerns have been raised about serious consequences of working long hours and excessive work. For example, death from overwork, or "*Karoshi*," has been estimated to account for 10% of all deaths among working Japanese males. As matters stand, claims of an increase

may only reflect the fact that more attention is devoted to the workaholism phenomenon, rather than an increase in its actual occurrence.

Bearing in mind the lack of agreement concerning the conceptualization and measurement of workaholism, prevalence estimates for workaholism consequently vary, are rather conflicting and not entirely clear. As workaholism is claimed to be a stable individual characteristic, it should be rather independent of organizational type or level of performance. According to the attraction–selection–attrition theory, it is still reasonable, to expect that workaholics are attracted to and may be overrepresented in certain work contexts and professions than in others (e.g. managerial and professional work versus clerical and blue-collar jobs). Furthermore, the limited number of studies aiming at estimating the prevalence of workaholism suggests no differences on the basis of background demographics such as marital status, race, or gender. Other researchers claim to have found higher scores on workaholism measures among men and younger workers than among women and older employees. All in all, it is premature to expect proper and well-documented estimates of prevalence rates of workaholism.

WORKAHOLISM AND ANTECEDENTS

In the following, the theoretical and empirical framework for potential antecedents of workaholism will be briefly presented, in addition to other central theories concerning antecedents of workaholism. Some authors have reported a link between workaholism and intrinsic and extrinsic job motivation. This may relate to innate basic psychological needs of autonomy, competence, and relatedness. According to the theory, these three motives are basic motivations for all human behaviors. The workaholic's behavioral pattern can be understood in terms of satisfying these basic needs. For example, when feeling incompetent, the person may work excessively in order to feel more competent, particularly if this motive is given high priority by the individual in question. Internal pressure or obsession with work could therefore be related to unsatisfied basic needs, which may be satisfied through work. According to this theory, need satisfaction is necessary for enjoyment of work. In line with this, Andreassen, Hetland, and Pallesen found that Drive was negatively related to autonomy and that Enjoyment of Work was positively related to autonomy and competence.

Reasoning along these lines, Csikszentmihalyi emphasized the concept of "flow" when describing an optimal work-experience, and it seems relevant to some workaholic types. "Flow-experiences" make individuals lose their sense of time and place and even forget biological needs when they are focused on a work task. Csikszentmihalyi suggested that goals, motivation, and intentions are central to achieving these good and optimal experiences. The theory of flow-experiences may reflect the proposed typologies of the goal-oriented workaholic, the achievement-oriented workaholic, and the enthusiastic workaholic.

Workaholism is considered by some to reflect one or several higher order personality traits. Theorists generally assume that such traits are relatively firmly established, differ between individuals, and affect behavior. It is also believed that most basic personality traits have a genetic basis and are further developed early in life. Three studies have linked workaholism with the Five Factor Model (FFM) of personality. The generic FFM of personality comprises Neuroticism, Extraversion, Openness to experience, Agreeableness, and Conscientiousness. Burke et al.'s study using WorkBAT reported that Drive was strongly and positively related to Neuroticism, whereas Work Involvement and Enjoyment of Work were positively related to Extraversion. Andreassen et al., also using the WorkBAT, found that Work Involvement was positively related to Neuroticism, Extraversion, Openness to experience, and Conscientiousness. Drive was positively related to Neuroticism and Conscientiousness, but negatively related to Agreeableness. Enjoyment of Work was positively related to Extraversion and negatively related to Agreeableness. Clark et al. using the WART found that Neuroticism was the only Big Five trait which was related (positively) to workaholism. Besides Neuroticism, workaholism was positively related to negative affectivity, narcissism, and perfectionism.

Some authors theoretically frame the structure, process, growth, and development of workaholism within addiction theory. Thus, these authors view workaholism as an addiction, comparable to other addictive behaviors such as alcoholism or pathological gambling. Addiction theory comprises two major explanatory models of addictive behavior: the medical and the psychological model. In short, the medical model of addiction emphasizes physical dependence on a substance (internal or external) accompanied by physical symptoms such as increased tolerance, withdrawal, and craving when the supply of the substance diminishes, is absent, or is removed. Based on this model, some authors have suggested that workaholic behavior is somatically stimulated by physiological activation produced by, for example, working against the clock to meet a deadline. In a similar vein, with reference to the psychological model of addiction, it has also been suggested that workaholism is developed and driven by a craving for approval, reward, and recognition rooted in narcissistic traits. Taken together, data supporting these assumptions have so far failed to provide empirical evidence that proneness to addiction is a causal factor in the development of workaholism.

Based on learning theory, the development of workaholism and corresponding behavior can be explained by various principles of learning. As long as the proper reinforcing conditions are present or absent, this model suggests that anyone can be led into or out of workaholism. Following the principles of operant reinforcement, workaholic behavior occurs, is nurtured, and sustained because similar behavior in the past has led to positive outcomes, such as praise and approval from peers, promotion, a good salary, or because the behavior has led to the avoidance of negative outcomes such as conflict at home, unwanted leisure, and socialization. In line with this, workaholism has previously been linked to salary increase and promotions. The principles of social learning can also explain how workaholism is molded. The individual may be influenced by observing the behavior of significant others, such as parents, peers, leaders, and by being exposed to models in the media. At present, however, the learning theory of workaholism remains unexplored.

Also cognitive perspectives on workaholism may add to our understanding of its antecedents. Basic cognitions such as schemata, core beliefs, expectations, attributions, and automatic thoughts are assumed to activate behavior. Thus, if a person holds self-schemata that he/she is a poor achiever, and holds a core belief that hard work equals success, then workaholic behavior may be activated as a result. Self-efficacy or the perceived ability to handle specific situations well may also be relevant in explaining development of workaholism. In general, people are motivated to enhance positive self-evaluations. Thus, when self-efficacy at work is greater than self-efficacy in non-work settings, it may drive the worker to choose work over non-work for the purpose of maintaining a positive self-image and in order to strengthen the sense of mastery. The cognitive perspective on workaholism has to some extent been tested empirically, but more empirical validation is needed. Burke et al.'s study using WorkBAT reported that all three workaholism components were positively related to generalized self-efficacy, with the strongest relation being to Enjoyment of Work.

Furthermore, some people have argued that workaholism should best be understood in terms of family-systems, with the family being the origin of the phenomenon. According to this perspective, workaholism is regarded as the result of certain family dynamics (e.g. over-responsibility, parentified children), affecting the individual within the system. In some cases, long working hours motivated by providing for one's family may reflect over-responsibility. In one study, it was found that students with high scores on the WART perceived their parents as more heavy workers than students with low scores on the WART. Still, as for the other theories discussed so far, the empirical basis for the family-system perspective is scarce.

Although the theories presented here have different explanations of workaholism, they should not be considered mutually exclusive. Workaholism is most probably formed by a variety of antecedents. In short, each theory may add a piece to the puzzle. Most recently, Ng et al. combined parts from several of these theories and proposed that workaholism is the combined result of dispositional traits (e.g. needs, values, traits), socio-cultural experiences (e.g. social learning, cultural emphasis on competence and competition), and behavioral reinforcements (e.g. organizational reward systems).

OUTCOMES AND CORRELATES OF WORKAHOLISM

Workaholism seems to entail both benefits and risks, and potential outcomes are better documented than the antecedents of workaholism. It is important to recognize, however, that outcome research has predominately relied on simple cross-sectional study designs. Both previous and recent research suggests that enthusiastic workaholic features (e.g. work enjoyment) are associated with higher scores on career satisfaction, work engagement, job satisfaction, and low intention to quit their jobs, compared to features characteristic of non-enthusiastic workaholics, who display the reverse relationships. Furthermore, research has found that workaholism in general is related to a broad range of work attitudes such as perfectionism, inability to delegate, and long working hours. This suggests that workaholism may be productive and profit the organization, but it may also result in destructive outcomes for the individual, team, and organization. So far no studies have been conducted on the work performance of workaholics using objective measures of the latter – underlining that spending long hours at work does not necessarily mean good and profitable performance.

Spending a great amount of time on work-related activities seems to be a core element in most definition of workaholism. Since time is a fixed unit, this must have consequences for time spent on non-work activities; thus workaholism may influence the domestic area. Extensive involvement in work, chronic stress, and time pressure, certain coping styles and personality characteristics, an unsupportive organizational climate and poor leadership, etc. have all been postulated as relevant predictors of work–family conflicts. Work–family conflicts are regarded as inter-role conflicts in which the role pressures from the work and family spheres are seen as being incompatible and are linked to negative physiological, behavioral, and psychological outcomes. Work–family conflicts and poor social functioning

outside work have also been proposed as potential consequences of workaholism. However, few studies have been conducted that examine the relationship between workaholism and the work–family interface. The limited number of such studies that have been conducted all conclude that workaholics report a higher degree of work–family conflict than non-workaholics. Bakker et al. using the DUWAS found that workaholism is related to reduced support provided to the partner through work–family conflict. In a recent study differentiating between positive and negative spillover, it was found that Drive was related to negative spillover between work and family, whereas, Enjoyment of Work was related to positive spillover between work and family. Work involvement showed inconsistent relationships with spillover. Although it has been suggested that the relationship between workaholism and work–life imbalance might be moderated by culture, a recent study of Caucasian and Black respondents revealed that cultural origin did not moderate this relationship.

It has further been suggested that workaholism is related to general life satisfaction. Again, since the workaholic's time is primarily devoted to work, how one experiences work will influence how life is experienced. Consequently, since enthusiastic workaholism is associated with passion and positive feelings about work, a large proportion of workaholics' time is spent on pleasurable (work) activities. Unsurprisingly, life satisfaction and purpose in life are reported to be high for these workaholics, whereas non-enthusiastic workaholic features have recently been empirically linked to the opposite pattern. In line with this, Andreassen, Hetland, Molde, and Pallesen using the WorkBAT found that Drive was negatively related to life satisfaction, whereas Work Enjoyment was positively related to life satisfaction. They also found that Drive was negatively associated with job satisfaction and that Enjoyment of Work was positively associated with job satisfaction. In one study using the DUWAS, workaholism was negatively correlated with perceived health and happiness. Chamberlin and Zhang found that student with high scores on the WART reported lower levels of psychological well-being, lower levels of self-acceptance and more physical health complaints compared to student with low scores on the WART. Another study based on the DUWAS showed that workaholism was positively correlated with ill-health and negatively associated with life satisfaction and self-reported job performance. Based on the WorkBAT, Andreassen et al. reported that Drive was positively related to subjective health complaints, while Enjoyment of Work was negatively related to health complaints. As sleep is a complex biological response it is highly vulnerable to stress. Two studies have specifically investigated how workaholism relates to sleep. In one study, subjects with the highest scores on the DUWAS had greater probability than workers with low scores to report insufficiency of sleep, sleepiness at work, and difficulties awakening in the morning as well as tiredness in the morning. In another study based on the WorkBAT, Andreassen et al. found that Enjoyment of Work was negatively related to insomnia whereas Drive was positively related to insomnia.

Generally, it has been argued that workaholism is detrimental to health. In fact, Ng et al. most recently argued that workaholism has both direct and indirect costs in the form of poorer physical and mental health because workaholics do not prioritize protective behavior such as leisure and exercise. The workaholic lifestyle may thus lead to increased blood pressure, cholesterol, poor sleep, etc. There are however no hard data to substantiate such claims, but associations between potential risk factors for illness and disease and workaholism have been postulated. Still, McMillan and O'Driscoll reported similarities in psychological, physical, and general health between workaholics and non-workaholics in their New Zealand study. Thus, more studies on the relationship between workaholism and parameters of somatic health are warranted.

Another potential and suggested consequence of workaholism is "burnout." Burnout is defined as a syndrome of emotional exhaustion, cynicism, and low professional efficacy. Based on the WorkBAT, Andreassen et al. found that Drive was positively related to burnout, whereas Enjoyment of Work showed the opposite relationship to this construct. Moreover, some studies have reported a link between workaholism and potential health risk related to "stress." Specifically these studies have shown that, when experiencing strain in their jobs, individuals scoring high on the Drive dimension or on Compulsive tendencies tend to report an increase in subjective stress-related somatic and psychological symptoms, whereas this response pattern has not been found in individuals with high scores on Enjoyment of Work. In one recent study using the DUWAS in a sample of junior doctors it was found that the relationship between workaholism on the one hand and well-being (job satisfaction, happiness, and perceived health) and burnout on the other was fully mediated by role conflicts.

As the stress response is more dependent on the perception of the stressor than the stressor in itself, the relationship between coping and workaholism should be investigated in future studies. In one study that included measures of coping, workaholism, and health, it was found that workaholism had a direct negative association with poor health (direct path). Workaholism was associated with better health through active coping (indirect path) and with poor health through emotional discharge (indirect path). Hence coping and emotional discharge partially seem to mediate the relationship between workaholism and health.

TREATMENT

Not much is actually known about treatment for workaholism as no controlled treatment/intervention study for workaholism so far has been conducted. Still, some suggestions and recommendations regarding therapy for this condition have been proposed as recently reviewed by Rebecca Burwell and Charles Chen. Inherent in the treatment perspective is the assumption that workaholism primarily represents a negative entity.

Several authors on workaholism have pointed to the fact that many workaholics typically deny their problem. When seen in therapy the workaholics often complain about marital problems and many therefore enters therapy with this focus (e.g. marital therapy).

Telling the workaholic to cut down on the number of working hours is normally warned against, as this is assumed to generate heavy therapeutic resistance. In terms of individual interventions, some argue for modifying the underlying personality factors that are assumed to be at play. In this regard some authors regard workaholism as a symptom of not being able to form authentic bonds to other people. In line with this, work is seen as a way to block out lack of intimacy. It has been suggested that therapy therefore should focus on the emotional consequences of not being able to create deep bonds to other people. One way of working with this may be to conduct cost–benefit analyses concerning the rewards experienced from work and the potential rewards which could be gained by creating deep emotional ties to others. Other authors have emphasized working with specific core beliefs about self, which many workaholics seem to be characterized by. Such core beliefs (e.g. "only when I work hard I will get the attention and love I deserve" and "work performance is the only way to prove my self-worth," etc.) can for instance be targeted and corrected through cognitive therapy. Others have focused more on changing the behavior of the workaholic more than his/her underlying psychological makeup. These approaches typically entail helping clients setting up boundaries between their private and their work arena, for instance, by the use of time management principles (planning how much time to be used at, for instance, work, to self, family, and play). Therapy may also involve identifying triggers (both internal in terms of dysphoric feelings and external in terms of specific situations) for work bouts. Another approach is the Workaholic Anonymous, which advocates a similar 12-step approach used for the treatment of other addictions. A potential effective treatment approach, either by itself or in combination with other approaches, is motivational interviewing (MI). MI represents techniques, including interviews and questions, where the clients themselves discover the negative aspects and consequences of their behavior. Focus is also upon how the clients want their life situation to be and what they can do in order to bring it there. Hence, MI aims at increasing and strengthening the client's motivation for change.

Interventions against workaholism may also comprise the organizational level. Managers are for instance recommended not to set unrealistic expectations for their followers. Managers should also be aware of the importance of themselves as role models, thus a workaholic manager may consequently be a bad example for his/her subordinates. Some companies arrange work–life balance programs in order to prevent and treat workaholism. It may also be important that managers are able to identify workaholic behavior and to encourage the person in question to prioritize recreation and his/her private life.

A newly adopted approach concerning the treatment of workaholism is based upon positive psychology. Within this orientation, focus is not on the shortcomings and problems inherent in human nature, but the emphasis is instead put on strengths and positive human qualities. Positive psychotherapy may involve techniques determining ones strengths and to use one strength in a new way every day for a week. Another technique is the gratitude letter technique where the client is asked to write a letter to someone they have a great gratitude to. The workaholic could for instance write and deliver such a letter to a person whom he or she has ignored due to his/her work investment. Another approach within this perspective is life quality therapy. The client is instructed to rate different life areas in terms of importance. Then the clients are to rate how much satisfaction is derived from each of these areas. The product (importance times satisfaction) yields a score reflecting life satisfaction. Many workaholics may by this method discover that their life quality is low. Quality of life therapy contains several therapeutic techniques aiming at increasing the life quality of the client. These techniques comprise self-care (e.g. exercise and rest), quality time (e.g. relaxing alone), and developing a guiding vision of what matters most in life. Specific interventions for every single life domain have also been developed. The main aim with life quality therapy is to have the workaholic to discover that happiness comes from many other sources than work, paving the way for cutting down on work and prioritizing other important life arenas.

FUTURE RESEARCH DIRECTIONS

The research on workaholism seems still to be in its infancy. Hence several important research questions

remain unanswered. First and foremost the field is in need of conceptual clarifications. The different measures of workaholism have, to a very limited extent, been cross-validated with each other, thus it is not known whether these actually measure the same construct(s). Related to this is the question pertaining to the subdimensions of workaholism. Is workaholism a multidimensional construct with both positive and negative aspects? Some argue, for example, that the term workaholism should be reserved for negative work-related attitudes, cognitions, and behaviors. The term "work addiction" may be a more suitable label for "negative workaholism." In addition, clear demarcations, both on a conceptual as well as on an empirical level, should be made between the term "workaholism" and related and perhaps overlapping concepts such as job satisfaction, job stress, passion toward work, work engagement, and work overcommitment. Consensus about the operationalization of workaholism would allow for comparisons of results across studies, development of norms as well as decisions concerning the use of cut-offs in order to make proper estimates of the prevalence of workaholism in different populations. Unfortunately, the situation in this field seems to suggest that it will take considerable time and effort to reach such a goal.

Future research should also investigate to what degree workaholism represents a rather stable individual characteristic and to what degree it does vary according to the context. As by far the great majority of studies of workaholism are cross-sectional, more studies using longitudinal designs are strongly warranted in order to be able to reveal some potential causal mechanisms concerning antecedents as well as consequences of workaholism. Most studies have also relied on single source of information (self-report only). Thus, studies collecting data from colleges (e.g. 360 degrees evaluation) as well as from marital partners and other family member would be highly welcomed. In addition, very few, if any, studies have linked workaholism to biological data/outcomes, such as immunological, hormonal, and metabolic parameters. Thus, the field is in demand of studies including assessment of such variables. As other non-chemical additions have been shown to be related to specific genotypes, future studies should also investigate whether this is the case for workaholism as well.

Although several therapeutic interventions for workaholism have been proposed, no randomized clinical trial of any workaholism treatment has ever been conducted. Hence, development and empirical investigations of interventions for workaholism should comprise another future research goal within this area. Future studies should also take cultural differences into consideration. Studies have, for example, shown that patterns of work and attitudes toward work differ between cultures in accordance to their emphasize on survival values, self-expression values, and mastery values. Gender differences concerning time devoted to work seem to be greater in masculine than in feminine societies; hence, this dimension should also be taken into consideration in future workaholism research.

SEE ALSO

Gambling, Video Game Addiction, Shopping Addiction, Exercise Dependence, Cognitive Factors in Addictive Processes, Emotions and Addictive Processes, Contextual Factors in Addiction, Epidemiology of Addiction, Families and Addiction, Personality and Addiction Processes, Stress and Addiction

List of Abbreviations

DUWAS	Dutch Workaholism Scale
FFM	Five Factor Model
SNAP	Schedule for Non-adaptive Personality Workaholism Scale
WART	Work Addiction Risk Test
WorkBAT	Workaholism Battery

Glossary

Agreeableness agreeableness is the tendency to be compassionate and cooperative rather than suspicious and antagonistic toward others.

Antecedent antecedent is a preceding occurrence, cause of event that existed or comes before something else; the antecedent of workaholism.

Autonomy autonomy is a noun and means self-governing. It is having freedom and independence as a person.

Big Five traits Big Five traits or factors (or Five Factor Model; FFM) of personality comprise five broad domains or dimensions of personality, which are used to describe personality in contemporary psychology: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism.

Burnout burnout is a psychological term for the experience of long-term exhaustion and diminished interest.

Cognition cognition is the scientific term for "the process of thought." Usage of the term varies in between disciplines; in psychology, and cognitive science, it usually refers to an information processing view of an individual's psychological functions.

Cognitive psychology cognitive psychology is a subdiscipline of psychology exploring internal mental processes and focuses on topics such as how people perceive, remember, think, speak, and solve problems.

Compulsive Tendency/Compulsive behavior compulsive tendency/compulsive behavior is a tendency or behavior motivated by factors that compel a person to act, sometimes against his or her own wishes.

Conscientiousness conscientiousness is the tendency to show self-discipline, act dutifully, and aiming for achievement; planned rather than spontaneous behavior.

Coping style coping style denotes the general approach in terms of cognitive, affective, and behavioral responses taken by the person for dealing with the anxiety triggering or challenging situations in life.

Correlation correlation is a mutual or reciprocal relationship between two or more variables, a statistic representing how closely two variables co-vary.

Drive drive is one dimension or aspect of workaholism as measured by the frequently used Workaholism Battery (WorkBAT) and reflects working due to feelings of being compelled or driven by inner forces. See Workaholism Battery.

Dysphoric feelings/Dysphoria dysphoric feelings/dysphoria is an unpleasant or uncomfortable mood, such as sadness (depressed mood), anxiety, irritability, or restlessness. Etymologically, it is the opposite of euphoria.

Dutch Work Addiction Scale (DUWAS) Dutch Work Addiction Scale (DUWAS) is a 10-item questionnaire to measure work addiction. The DUWAS comprises two dimensions: Working Excessively and Working Compulsively.

Enthusiastic workaholic enthusiastic workaholic is a workaholic who is characterized by high levels of Work Involvement, Drive, and Enjoyment of Work as measured by the Workaholism Battery (WorkBAT). See Workaholism Battery.

Enjoyment of Work Enjoyment of Work is one dimension or aspect of workaholism as measured by the Workaholism Battery (WorkBAT). See Workaholism Battery.

Extraversion extraversion is the tendency to have energy, positive emotions, surgency, and the tendency to seek stimulation in the company of others.

Extrinsic job motivation extrinsic job motivation comes into play when a person is motivated to work primarily due to factors external to him or her (like money or promotion).

Five Factor Model (FFM) see Big Five traits.

Flow flow is a mental state originally described by Mihály Csikszentmihalyi in which a person becomes fully immersed in a feeling of energized focus, full involvement, and success in the process of the activity.

Intrinsic job motivation intrinsic job motivation occurs when people are internally motivated to work because the work or activity in itself either brings them pleasure, an experience of importance or provides a significant learning opportunity.

Likert scale likert scale is a scale commonly used for response alternatives in surveys in which the respondents specify their level of agreement to a statement by endorsing one of normally five or seven response alternatives. The middle alternative is most often neutral, whereas the others reflect some level of agreement or disagreement with the statement. The scale is named after its inventor, psychologist Rensis Likert.

Motivational interviewing (MI) motivational interviewing (MI) refers to a counseling approach in part developed by the clinical psychologists William R. Miller, PhD and Stephen Rollnick, PhD. It is a client-centered, semi-directive method of engaging intrinsic motivation to change behavior by developing discrepancy and exploring and resolving ambivalence within the client.

Narcissism narcissism trait reflects a general tendency of egotism, vanity, simple selfishness, envy, using and seeing others as objects and self-grandiosity.

Neuroticism neuroticism is the tendency to easily experience unpleasant emotions, such as anger, anxiety, depression, or fear.

Non-enthusiastic workaholic non-enthusiastic workaholic is a workaholic who is characterized by high levels of Work Involvement, Drive, and low Enjoyment of Work. See WorkBAT.

Openness to experience openness to experience is the tendency to appreciate art, emotions, adventure, unusual ideas, and variety of experience.

Operant conditioning/Reinforcement operant conditioning/reinforcement is a form of learning where an individual modifies the occurrence and form of its own behavior due to learning associations between the behavior and its consequence. The consequences are denoted as reinforcement and punishment. Reinforcement increases the frequency of the behavior and is either positive (a positive stimulus occurs following the behavior) or negative (an aversive stimulus disappears following the behavior). Punishment decreases the frequency of the behavior and is either of type I (introduction of an aversive stimulus) or type II, also called penalty (removement of a positive stimulus following the behavior).

Observational learning/Social learning observational learning/social learning is learning that occurs as a function of observing, retaining, and replicating behavior observed in ones environment or other people.

Perfectionism perfectionism is a personality trait manifested by the rejection of personal achievements falling short of perfection, often leading to distress and self-condemnation. It is a disposition to feel that anything less than perfect is unacceptable.

Personality traits personality traits are relatively firmly established general ways of responding, thinking, and feeling and which differ between individuals. Trait theory suggests that basic personality traits have a genetic basis and are further developed early in life.

Positive psychology positive psychology is a recent branch of psychology in which emphasis is put on factors and processes relevant for growth, positive development, achievement, and success.

Psychometric properties psychometric properties are characteristics of tests and other measures of human behaviors, thoughts, or feelings that describe their basic measurement features. Most common psychometric properties express some aspects of the stability or consistency of the instrument (denoted as its reliability) and some aspects indicating whether the instrument measures what it is intended to measure (denoted as its validity).

Schedule for Non-adaptive Personality Workaholism Scale (SNAP) Schedule for Non-adaptive Personality Workaholism Scale (SNAP) is an 18-item questionnaire developed by Lee Anna Clark in 1993 to measure workaholism. The questionnaire largely measures obsessive–compulsive traits.

Self-efficacy self-efficacy is defined in a variety of ways: as the belief that one is capable of performing in a specific manner to attain specific goals, as a person's general belief about their capabilities to produce designated levels of performance that exercise influence over events that affect their lives.

Spillover spillover is the transfer of mood, energy, and skills from one sphere to another. Spillover can be positive as well as negative (e.g. work–family facilitation and work–family conflict) and can operate in both directions (e.g. work-to-family and family-to-work).

Subjective health complaints subjective health complaints are health complaints with few or no objective findings, and which typically include muscle pain, tiredness, mood disturbances, fatigue, headaches, sleep problems, and gastrointestinal complaints.

Twelve-step program twelve-step program is a set of 12 guiding principles outlining a course of action for recovery from addiction, compulsion, or other behavioral problems. The program is normally run by former addicts forming a fellowship.

Work Addiction Risk Test (WART) Work Addiction Risk Test (WART) is a 25-item questionnaire developed by Robinson (1998) to measure work addiction. The questionnaire measures five dimensions of workaholism: Compulsive Tendencies, Control, Impaired Communication/Self-Absorption, Inability to Delegate, and Self-Worth.

Work Engagement work engagement is characterized by energy, involvement, and efficiency, and as a relatively stable condition characterized by vigor, dedication, and absorption.

Work–family conflicts work–family conflicts are inter-role conflicts in which the role pressures from the work and the family sphere are perceived as incompatible.

Work Involvement Work Involvement is one dimension or aspect of workaholism as measured by the frequently used Workaholism Battery (WorkBAT). See WorkBAT.

Workaholic workaholic or work addict is a person obsessively addicted to work, person with a compulsive need to work.

Workaholic Anonymous Workaholic Anonymous is a fellowship of individuals who share their experience, strength, and hope with each other that may solve their common problems and help others to recover from workaholism.

Workaholism workaholism or work addiction is compulsiveness about working and/or working excessively.

Workaholism Battery (WorkBAT) Workaholism Battery (WorkBAT) is a 25-item questionnaire developed by Spence and Robbins (1992) to measure workaholism. The questionnaire measures three dimensions of workaholism: "Work Involvement" (eight items), reflecting the need to spend time efficiently both at work and when off work, blurred boundaries between work and private life and the inability to relax. "Drive" (seven items) reflects internal motivation for work and the frequency of thinking about work. "Enjoyment of Work" (10 items) assesses satisfaction from work.

Further Reading

Andreassen, C.S., Griffiths, M.D., Hetland, J., Pallesen, S., 2012. Development of a work addiction scale. Scandinavian Journal of Psychology 53, 265–272.

Andreassen, C.S., Hetland, J., Pallesen, S., 2010. The relationship between workaholism, basic needs satisfaction at work and personality. European Journal of Personality 24, 3–17.

Andreassen, C.S., Hetland, J., Pallesen, S. Workaholism and work–family spillover in a cross-occupational sample. European Journal of Work and Organizational Psychology, DOI: 10.1080/1359432x.2011.626201.

Andreassen, C.S., Ursin, H., Eriksen, H.R., 2007. The relationship between strong motivation to work, "workaholism", and health. Psychology and Health 22, 615–629.

Andreassen, C.S., Ursin, H., Eriksen, R.H., Pallesen, S., 2012. The relationship of narcissism with workaholism, work engagement, and professional position 40, 881–890.

Bonebright, C.A., Clay, D.L., Ankenmann, R.D., 2000. The relationship of workaholism with work-life conflict, life satisfaction, and purpose in life. Journal of Counseling Psychology 47, 469–477.

Burke, R.J., 2000. Workaholim in organizations: concepts, results and future research directions. International Journal of Management Reviews 2, 1–16.

Griffiths, M.D., 2011. Workaholism: A 21st century addiction. The Psychologist: Bulletin of the British Psychological Society 24, 740–744.

Kanai, A., Wakabayashi, M., Fling, S., 1996. Workaholism among employees in Japanese corporations: an examination based on the Japanese version of the workaholism scales. Japanese Psychological Research 38, 192–203.

Killinger, B., 1991. Workaholics: The Respectable Addicts. Firefly, Buffalo, NY.

McMillan, L.H.W., O'Driscoll, M.P., Burke, R.J., 2003. Workaholism: a review of theory, research and future directions. In: Cooper, C.L., Robertson, I.T. (Eds.), International Review of Industrial and Organizational Psychology, vol. 18. Wiley, New York, pp. 167–189.

Ng, T.W.H., Sorensen, K.L., Feldman, D.C., 2007. Dimensions, antecedents, and consequences of workaholism: a conceptual integration and extension. Journal of Organizational Behavior 28, 111–136.

Robinson, B.E., 1998. Chained to the Desk: A Guidebook for Workaholics, Their Partners and Children, and the Clinicians Who Treat Them. New York University Press, New York.

Schaufeli, W.B., Taris, T.W., Van Rhenen, W., 2008. Workaholism, burnout, and work engagement: Three of a kind or three different kinds of employee well-being? Applied Psychology: An International Review 57, 173–203.

Scott, K.S., Moore, K.S., Miceli, M.P., 1997. An exploration of the meaning and consequences of workaholism. Human Relations 50, 287–314.

Shimazu, A., Schaufeli, W.B., Taris, T.W., 2010. How does workaholism affect worker health and performance? The mediating role of coping. International Journal of Behavioral Medicine 17, 154–160.

Spence, J.T., Robbins, A.S., 1992. Workaholism – definition, measurement, and preliminary results. Journal of Personality Assessment 58, 160–178.

Taris, T.W., Schaufeli, W.B., Verhoeven, L.C., 2005. Workaholism in the Netherlands: measurement and implications for job strain and work–nonwork conflict. Applied Psychology – An International Review-Psychologie Appliquee – Revue Internationale 54, 37–60.

CHAPTER

85

Shopping Addiction

*Michel Lejoyeux**, *Aviv Weinstein*[$]

[*]Paris 7 University, Paris, France [$]Hadassah Medical Organization, Jerusalem, Israel

OUTLINE

INTRODUCTION AND DEFINITION

Shopping addiction is a frequent and often underrecognized form of behavioral addiction. For shopping addicts, buying becomes uncontrolled and repetitive and leads to severe financial and psychological consequences. Addicts do not buy only items they need or they like. They also experiment craving for buying clothes or other items that they do not use after their purchase. They really need to spend their money and are anxious to miss a good opportunity to buy something. This form of addiction represents a pathological form of a normal behavior, stimulated by advertising and considered as a marker of individual and collective health. First descriptions of shopping addiction appeared in the Roman law to condemn spendthrifts who refused or rapidly spent the legacy of their parents. They refused at the same time transmission of family values and brought their own family about the ruin.

Emile Kraepelin, a century ago, proposed the term oniomania (from the Greek "onios" or "for sale"). According to Eugen Bleuler, shopping addiction could be considered as a type of monomania or instinctive impulse. Monomania was also described as an excessive focusing of attention on one object or issue. In 1960, "prodigality" was added to a French Manual of Psychiatry, defined as an abnormality of the conservation instinct, affecting object properties. More recent and consensual definitions describe shopping addiction as a repetition of excessive and impulsive buying entailing financial and familial problems.

Susan McElroy et al. in 1994 specified benchmarks for shopping in its compulsive dimension. These criteria include many clinical characteristics of addiction. They describe equivalents of craving and withdrawal applied

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00085-1

to shopping behavior. They also exclude buying related to a manic state.

Inappropriate preoccupations with buying or shopping, or inappropriate buying or shopping impulses or behavior, as indicated by at least one of the following:

- Frequent preoccupations with buying or impulses to buy that are experienced as irresistible, intrusive, and/or senseless.
- Frequent buying of more than can be afforded, frequent buying of items that are not needed, or shopping for longer periods of time than intended.
- The buying preoccupations, impulses or behaviors cause marked distress, are time-consuming, significantly interfere with social or occupational functioning, or result in financial problems (e.g. indebtedness or bankruptcy).
- The excessive buying or shopping behavior does not occur exclusively during periods of hypomania or mania.

Other definitions distinguish shopping addiction from normal economic temptations. Impulse shopping is a classical aspect of the consumer's lifestyle and a focal point of considerable marketing management activity. Economic studies found that 27–62% of department store purchases fall into the impulse category. One-third of the people, in the general population, experience shopping impulses. Marketing methods such as development of credit cards, cash machines, instant credit, home shopping networks, and telemarketing make it easier than ever before for consumers to purchase items on impulse. In the modern marketplace, spontaneous urges to buy and consume are normal phenomenon. Everyday consumers receive, via the web or in the real life, these stimulations and need to cope with the impulse to consume, "to have it all now." Bargains and new methods of marketing induce the shopping "spree." Confronted by the array of possibilities, each subject modulates his or her own behavior according to their particular level of impulsivity and frugality. Development of online buying has still increased the risk of shopping addiction. The Internet is becoming a basic feature of global civilization. Informative, convenient, and entertaining, the Internet has changed the ways people shop, work, and spend their leisure time.

Whatever shopping occurs in real life or online, it contains a part of emotion. Valence et al. distinguished four types of "pathological consumers:"

1. The emotional reactive consumer places some importance on the symbolism of the product, frequently presenting compensatory consumption and emotional motives.
2. The impulsive consumer feels a sudden spontaneous desire to buy, living in a state of probable psychological disequilibrium with a conflictual struggle between the id and the superego.
3. The fanatical consumer: He is often interested in only one product and is consumed by a monomania (e.g. either records, books, or clothing). This person displays an enthusiasm and an intense devotion.
4. The compulsive consumer tries to reduce psychological tension and sees the act of buying as a means to reduce this tension or anxiety. He or she often does not seek the possession of goods, but rather the immediate reduction of a state of tension.

According to this classification, shopping addiction concerns subjects who are simultaneously emotional reactive, impulsive, and compulsive consumers. Such subjects are especially sensitive to marketing and social stimulation. Most patients presenting shopping addiction reported that having access to credit cards, money, shopping catalog, online websites or home shopping television programs triggered or increased their compulsive buying.

The two best distinctions between normal urges to buy and shopping addiction are the negative consequences of the behavior and the fact that items bought compulsively are not used as much as expected. Shopping addiction buying must indeed be considered as pathological since it causes personal distress, consequential financial debts, and marital and family disruption. While initially providing some perceived benefits, pathological shopping addiction typically becomes very difficult to stop. It ultimately results in harmful consequences like unmanageable indebtedness or bankruptcy.

EPIDEMIOLOGY

Shopping addiction is primarily a female problem. The typical compulsive buyer is a 36-year-old educated woman. The mean age of onset is 30 years. Clinical studies suggest a beginning of the compulsive behavior as soon as people reach financial independence or earn their first wage. Population-based surveys in the United States show that up to 5% of adults may have shopping addiction. Five million Americans spend excessively. Compulsive buying has significantly risen in developed economies and with the web. A study of 200 women who shopped at a renowned Parisian department store found that 32.5% met criteria for shopping addiction. There was no difference from a control group in age, education level, or professional status. Among 1490 customers of an Internet women's clothing retailer, 17.7% were shopping addicts. In a population of 203 medical students, 11% were shopping addicts.

ADDICTIVE NATURE OF THE DISORDER

Shopping addiction can be included in the "addictive spectrum" due to its clinical characteristics in common with classical addictive disorders. Shopping addicts experiment all signs characteristics of dependence and especially craving, loss of control and social and personal consequences of shopping. *But there are also differences from other behavioral addictions – particularly regarding pleasure or positive reinforcement.*

Intoxication

Most patients describe their shopping experiences as "a high," "a buzz,""a rush." They experience a positive feeling when they buy. Their sensation can be compared with the effect of alcohol or drug intoxication.

Craving

Shopping addicts experience irresistible urges, uncontrollable needs, and mounting tension that could only be relieved by buying. The tension is temporarily assuaged by the purchase but a feeling of guilt quickly replaces the euphoria associated with the purchase. A stimulus (either an internal stimulus, such as emotion, or an external one, such as advertising, stay in a shop, fear of missing a good shopping opportunity) can trigger a thought, which in turn creates a craving. This craving can be strengthened by a permissive thought which leads to an excessive purchase. A person can suffer from a sad mood, in a specific moment. In his mind could appear the idea that he will be less sad if he or she could please themselves with a present. She becomes convicted that buying something will make her happier. A craving for buying something (whatever it is) appears. She has on her mind some "good reasons" to do that, such as "I am worth it"… At the end, the patient enters a shop and buys compulsively items she does not really need. Dependents of online shopping experience craving when they see advertising on their computer screen or when they receive mails suggesting them good affairs to make.

Loss of Control

Shopping addicts consider their purchases as occasions not to be missed. They are often disappointed afterward and become angry when they see the amount of money spent. Compulsive buyers would buy clothes from famous designers and top of the line items in order to impress other people. Loss of control involves in the number of items bought (shoes from all colors available in the shop) or the price unaffordable for them. They also shop online more often than controls do. Buying is also impulsive: the duration between desire to buy and purchase is shorter than for normal buying. Due to this loss of control, purchasing is often a lonely activity. Patients are ashamed of their uncontrolled behaviors and do not want to be seen in a situation of pathological consumption.

Equivalent of Withdrawal

Shopping addicts consider that they miss important occasions to buy items when they cannot go shopping and they regret for the items they did not buy. They feel anxiety and tense comparable to manifestations of withdrawal.

Negative Consequences

Like other addictions, addictive shopping can result in a number of difficulties, including substantial debt, legal problems, personal distress, and marital conflict. More than half of shopping addicts have debts and are unable to make payments. They suffer remarks and reproaches from relatives and friends and more rarely they face criminal or legal proceedings. In most cases, what they spend exceeds what they can afford. Shopping addicts take out loans or revolving credit and soon find themselves trapped in a vicious cycle of debt. They borrow large amounts of money from their friends or family. They make promises to pay back what they know that they cannot fulfill. In some extreme cases, they steal to continue shopping. Friends, relatives, and co-workers don't understand their passion for shopping. They start distrusting them. Shopping addicts become rejected and alienated because of their behavior. They are accused of being irresponsible and they hide what they buy. Relatives are suspicious or overly protective, destroying their credit cards or putting them on an allowance.

OTHER CLINICAL ASPECTS

The nature of products that patients buy during compulsive buying episodes are clothing (95.8%), shoes (75%), jewelry (41.7%), cosmetics (33.3%), antiques (25%), records, CDs (20.8%), cars (16.7%), household items (12.5%), and books (12.5%). Men often buy larger items (e.g. furniture, computers, and stereo equipment) in addition to clothing. Websites recently modified the type of items bought in an addictive manner. Items are often used to enhance both others and one's own image of self. In almost every case, the articles are not bought because they are needed or are a bargain, or even out

of an intrinsic desire for the thing itself. Shopping addicts report that about half of the items purchased during uncontrolled buying episodes were used minimally or not at all. Many patients hoard the items they bought and some gave them away as gifts.

Two basic dimensions are found among shopping addicts: tendency to spend and post-purchase guilt. Tendency to spend corresponds to the six following affirmations:

When I have money, I cannot help but spend part or all of it.

I am often impulsive in my buying behavior.

As soon as I enter a shopping center, I have an irresistible urge to go into a shop and buy something.

I am one of those people who often respond to direct mail offers.

I have often bought a product that I did not need, while knowing that I had very little money left.

I am a spendthrift.

Post-purchase guilt is assessed by the two following items:

At times, I have felt somewhat guilty after buying a product, because it seemed unreasonable.

There are some things I buy that I do not show to anybody for fear of being perceived as irrational in my buying behavior.

The majority of patients describe the course of their addiction as chronic. More rarely, they report that their shopping addiction is episodic, with periods of free of buying symptoms. Patients typically report a mean peak frequency of buying episodes of 17 months with episodes lasting from 1 to 7 h in duration.

Shopping addicts don't seek bargains. They consider their purchases as occasions not to be missed and are often disappointed afterward. They can buy clothes from famous designers and top of the line items in order to impress other people. They also shop online more than controls do. Most often, purchasing is a lonely and highly pleasurable activity. Shopping addicts also tend to be very materialistic individuals, to have a low self-esteem and to associate social status with the activity of buying. They buy items more often to impress others and tend to consider their purchases as personally gratifying. Designer's brands induce a great motivation to spend money. Status consumption requires them to continually increase their conspicuous signals of wealth and power. Lastly, shopping addicts are regularly deceived by items they buy. They use them less than expected, try to give them back to shops and sometimes hide them or throw them away. There are also anecdotal cases of compulsive buyers who do not enjoy shopping but enjoy the returning of goods. This is followed by a feeling of empowerment – "the product is not good enough" and it is followed by the return of cash which is spent on buying new goods but not for paying the credit card. Also, compulsive buying often runs in the family and can be used as mother–daughter bonding.

Among a population of medical students, we found that shopping addicts made more often self-gifts. They did not buy more often during bargain periods and did not consider more often their purchases as exceptional or as occasions not to be missed. Percentage of purchases less used than expected was higher in shopping addicts. Their purchases more often increased their level of self-esteem.

We also identified a different relation to money among shopping addicts. The Yamauchi and Templer's money attitude scale (MAS) identifies five factors describing the relation to money. *Power-Prestige* is a tendency to use money to influence and impress others. *Retention Time* assesses the tendency of careful financial planning and *Distrust* represents hesitant, suspicious, and doubtful attitude toward situations involving money. *Anxiety* measures to what extent money can be seen as a source of anxiety and *Affair missing* reflects fear of missing a good opportunity to buy an item. The MAS scores were higher in the shopping addict group for two subscales: *Distrust* and *Bargain missing*. Correlation studies (Pearson test) showed that severity of shopping addiction is significantly correlated to three subscales *Distrust*, *Anxiety*, and *Affair missing*.

We confirmed the diagnostic of shopping addiction with the Questionnaire about Buying Behavior (QABB) which consists of 19 *yes–no* items representing major basic features of compulsive buying (e.g. urges to shop and buy, negative feedback from family and friends, post-purchase guilt) based on the McElroy et al. criteria. The QABB is usually self-administered. This questionnaire, useful for clinical research, cannot replace clinical evaluation. Clinicians must inquire in detail about buying attitudes, the extent of preoccupation with buying and shopping, urges to buy, feeling and thoughts associated with buying, and the interference with social, financial, and occupational functioning.

DIFFERENTIAL DIAGNOSIS

Normal Buying Impulses

Shopping addiction must be distinguished from normal buying impulse. These urges lead to unplanned purchases which correspond to the difference between a consumer's total purchases at the completion of a shopping trip, and those that were listed as intended purchases prior to entering a store. Impulse buying is a pervasive and distinctive aspect of the consumer's lifestyle and a focal point of considerable marketing management activity. One-third of the people in the general population experience buying impulses.

Marketing methods such as development of credit cards, cash machines, instant credit, home shopping networks, and telemarketing make it easier than ever before for consumers to purchase items on impulse. In the modern marketplace, spontaneous urges to buy and consume are thus normal daily phenomena. Everyday consumers receive these stimulations and need to cope with the impulse to consume, "to have it all now."

Collectors

Collectors and shopping addicts both spend a lot of time, money, and energy in the hunt for "precious" goods but the fate of the items and the evolution of the behavior are different. Collectors show their goods with pride once acquired. They often display them, in a glass case, for example. Shopping addicts don't have such a passionate relationship with their goods, which often lose their attraction after purchase. Goods are often hidden in a closed place. For collectors, purchase is only a way to collect a group of objects. The pleasure is in the possession, more than the acquisition. For shopping addicts, the object of the addiction is the shopping itself.

Hoarding

Compulsive hoarding has been defined as the acquisition of, and failure to discard, possessions which appear to be useless or of limited value. People hoard many kinds of items like "free things" ("freebies"), as well as things other people have discarded. Free newspapers and unclaimed magazines are often kept by compulsive hoarders. Association between shopping addiction and hoarding is possible.

Mania

Unconsidered buying are common symptoms of a maniac episode. The disordered behavior comes from a loss of inhibition and the feeling of being almighty. In maniac crises patients make excessive purchases in an impulsive and euphoric way and lose all sense of reality. Diagnostic criteria for shopping addiction exclude buying sprees related to mania.

ADDICTIVE AND PSYCHIATRIC COMORBIDITY

Addiction is often associated with "classical" or chemical dependence on alcohol and nicotine. Among a population of 500 female patients consulting their general practitioner, shopping addiction was present in 97 cases (38%) and more frequent among patients with nicotine dependence (58 cases, 45%) than among those who were not dependent on nicotine (39 cases, 31%). These results suggest that a common addictive spectrum could expose to both addictions.

Shopping addiction is often *associated with depression and/or low self-esteem*. Specific studies found a prevalence rate of 31.9% for shopping addiction among individuals hospitalized for depression. When buyers are asked to describe how they feel before a typical buying episode, 53% say they are sad or depressed and 21% tense or anxious. When asked about their feeling after a buying episode, 42% say they feel guilty and 21% sad or depressed. Shopping addicts are more likely to engage in their pathological behavior when they experience negative emotions. They obtain a temporary relief from their negative emotions and decreased self-esteem. A positive and significant relationship was also found between anxiety and shopping addiction. Shopping addicts use the buying activity as a means of relieving stress and low self-esteem.

Assessment of 1500 consecutive general medical outpatients in Paris identified 60 patients with obsessive–compulsive disorder (OCD), of whom 14 (23%) also had shopping addiction. The prevalence of shopping addiction among patients without OCD was 0.3%. Patients with OCD and shopping addiction had higher scores on the CAGE questionnaire (assessing alcoholism).

Shopping addiction and eating disorders frequently co-occur. One study found that 20% of compulsive buyers also had an eating disorder; conversely, 17.6% of binge eaters had shopping addiction. Among a group of patients with a variety of eating disorders, the lifetime prevalence of shopping addiction was 11.8%.

Shopping addiction can also be a part of the compulsive–impulsive spectrum. Even if some patients of shopping addiction are impulsive buyers, the obsessive–compulsive aspect is often prominent. Also, it looks like shopping addiction is associated with negative affect and its relief and not with positive reinforcement. This is different from other addictions which people also find pleasurable (drugs, alcohol, videogames, etc.).

TREATMENT

The most effective part of the treatment is cognitive behavioral therapy. Pharmacological agents are limited to patients presenting another psychiatric disorder (anxiety or depression). Few medications showed an anti-addictive effect on shopping behavior. If shopping addiction is associated with OCD and addiction, it might be due to low serotonin levels and therefore Selective Serotonin Reuptake Inhibitors (SSRIs) would be relevant

whereas when addiction is pleasurable (e.g. for psychostimulants) there is a role for dopamine in the addiction and possibly in treatment. This hypothesis, however, needs to be confirmed by further experiments.

Cognitive Behavioral Therapy

The cognitive part of the therapy consists in identifying, working on, and changing thoughts and feelings related to buying. Patients often have difficulties identifying their feelings. The therapist should encourage them to use a daily self-monitoring diary of their craving for shopping in general and/or episodes of addictive shopping. Patients are asked to keep track of the kind of items purchased, the time spent in addictive shopping, and the amount of money spent if they failed to resist. During the sessions the therapist analyzes the patient's relationship to pleasure to buy and to money. Dysfunctional beliefs about shopping and the power that derives from it are identified and corrected. Step by step, patients admit that the euphoria they feel when buying a new item is an illusion. A functional analysis shows to them that their dysfunctional beliefs result from personal history and education background.

Treatment for compulsive buying tends to focus on teaching patients alternative responses to cope with negative emotions. The therapist trains the patient to solve his problems without the help of addictive shopping. He finds other behaviors which are less impulsive and more constructive. Therapist may propose homework assignments from one session to another in order to build on that training of alternative behaviors. Homework can also consist in exposition experiments. One such assignment might be to go to the supermarket or another shop, not buy anything and then leave the shop. Shopping addicts are invited to find other sources of pleasure and to rediscover what activities they consider pleasant or effective to release tension. Situations at high risk of relapses must be identified by the patient (e.g. going to a clothes' shop after a stressful situation, or entering into a store without a prepared list of needed items). The more the patient learns to anticipate such risks, the more efficient he will be when exposed to risk of relapse. Cognitive therapy should set up a rescue plan that indicates what to do when the patient faces up to a situation of extreme craving. It must be set up before any crisis so that the patient knows what to do when he is in a risk situation.

Group Therapy

Like most addictive disorders, shopping addiction represents a good indication for group therapy. Dynamic of the group reduces feeling of shame and denial reactions. Patients exchange around a theme (always including a shopping problem). They feel understood by other people suffering from the same addiction. Group sessions reinforce individual therapy and the patient's will to change. It also allows positive identification to shopping addicts who are no more dependents. A few support groups for individuals dealing with shopping addiction exist in different parts of the world. They help patients to find hope and motivation to cope with their craving and their social and financial difficulties.

Pharmacotherapy

Few controlled studies have assessed the effects of pharmacological treatment on shopping addiction. Most of them involved selective serotonin inhibitors. Fluvoxamine did not appear superior to placebo. Citalopram demonstrated a weak efficiency on a limited number of patients. Small open-label studies with other antidepressants (often in combination with mood stabilizers), or the opioid antagonist naltrexone reported improvement in a few number of patients. Clinical practice shows that antidepressants improve both depression and shopping addiction. No study, however, has been specifically conducted on this subgroup of patients presenting shopping addiction and depression. Globally, it cannot be proven that pharmacological treatment of shopping addiction compulsive buying is more effective than placebo.

CONCLUSION

Shopping addiction is a behavioral addiction that may be increasing in prevalence because of shopping availability especially via the Internet. The main characteristic of addictive buying is that items bought are rarely used. They consider their purchases as occasions not to be missed and are often disappointed afterward. Diagnostic criteria for shopping addiction parallel those for substance addictions, pathological gambling, or other impulse control disorders. Much future research is needed to improve the understanding of shopping addiction and to standardize regularly effective psychotherapies and pharmacological treatments.

In addiction, it looks like shopping addiction leans toward the obsessive compulsive end of the impulsive–compulsive spectrum. Secondly, shopping addictions is more associated with relief of depression and anxiety rather than pleasure or positive reinforcement. This has clinical implication for treatment with SSRIs. Thirdly, there are diagnostic issues whether to classify shopping addiction together with other behavioral

addictions (e.g. other behavioral addictions such as Internet and videogame addiction are found to be not only pleasurable but relief from negative affect). It is debated whether shopping addiction should be classified as an impulse control disorder. Its comorbidity with depression might be an obstacle for its inclusion as a separate diagnostic category. Due to these considerations, shopping addiction might be problematic for inclusion in the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-V).

Fourthly, very little is known about the psychobiological or pharmacological mechanisms underlying shopping addiction, there are hardly any genetic studies or brain imaging that might explain the neuropharmacological basis of this addiction. All these issues are important when considering diagnosis, mechanism of action, and treatment of shopping addiction.

List of Abbreviations

MAS money attitude scale
OCD obsessive–compulsive disorder
QABB Questionnaire about Buying Behavior

Further Reading

Adam, P., Richoux, C., Lejoyeux, M., 2008. Screening for impulse control disorders among patients admitted to a French emergency service. The Open Psychiatry Journal 2, 30–36.

Billieux, J., Rochat, L., My Lien Rebetz, M., Van Der Linden, M., 2008. Are all facets of impulsivity related to self-reported compulsive buying behavior? Personality and Individual Differences 44, 1432–1442.

Black, D.W., 2007. A review of compulsive buying disorder. World Psychiatry 6, 14–18.

Black, D.W., Repertinger, S., Gaffney, G.R., Gabel, J., 1998. Family history and psychiatric comorbidity in persons with compulsive buying: preliminary findings. American Journal of Psychiatry 15, 960–963.

Christenson, G., Faber, R.J., DeZwaan, M., Raymond, N.C., Specker, S.M., Ekern, M.D., Mackenzie, T.B., Crosby, R.D., Mussel, M.P., Mitchell, J.E., 1994. Compulsive buying: descriptive characteristics and psychiatric comorbidity. Journal of Clinical Psychiatry 55, 5–11.

Dell'Osso, B., Altamura, C., Allen, A., Marazziti, D., Hollander, E., 2006. Epidemiological and clinical updates on impulse control disorders: a critical review. European Archives of Psychiatry and Clinical Neuroscience 256, 464–475.

Faber, R.J., O'Guinn, T.C., 1989. Classifying compulsive consumers: advances in the development of a diagnostic tool. In: Srull, T.K. (Ed.), Advances in Consumer Research, vol. 16. Association for Consumer Research, Provo, UT, pp. 738–744.

Goodman, A., 2008. Neurobiology of addiction: an integrative review. Biochemical Pharmacology 75, 266–322.

Grant, J.E., 2003. Three cases of compulsive buying treated with naltrexone. International Journal of Psychiatry in Clinical Practice 7, 223–225.

Kim, S.W., 1998. Opioid antagonists in the treatment of impulse-control disorders. Journal of Clinical Psychiatry 59, 159–164.

Koran, L.M., Faber, R.J., Aboujaoude, M.A., Large, M.D., Serpe, R.T., 2006. Estimated prevalence of compulsive buying behavior in the united states. American Journal of Psychiatry 163, 1806–1812.

Kukar-Kinney, M., Ridgway, N.M., Monroe, K.B., 2009. The relationship between consumers' tendencies to buy compulsively and their motivation to shop and buy on the Internet. Journal of Retailing 85, 298–307.

Krueger, D.W., 1988. On compulsive shopping and spending: a psychodynamic inquiry. American Journal of Psychotherapy 42, 574–584.

Lejoyeux, M., Adès, J., Tassain, V., Solomon, J., 1996. Phenomenology and psychopathology of uncontrolled buying. American Journal of Psychiatry 155, 1524–1529.

Lejoyeux, M., Bailly, F., Moula, H., Loi, S., Ades, J., 2005. Study of compulsive buying in patients presenting obsessive-compulsive disorder. Comprehensive Psychiatry 46, 105–110.

Lejoyeux, M., Haberman, N., Solomon, J., Ad'es, J., 1999. Comparison of buying behavior in depressed patients presenting with or without compulsive buying. Comprehensive Psychiatry 40, 51–56.

Lejoyeux, M., Kerner, L., Thauvin, I., Loi, S., 2006. Study of impulse control disorders among women presenting nicotine dependence. International Journal of Psychiatry in Clinical Practice 10, 241–246.

Lejoyeux, M., Mathieu, K., Embouazza, H., Huet, F., Lequen, V., 2007. Prevalence of compulsive buying among customers of a Parisian general store. Comprehensive Psychiatry 48, 42–46.

Lejoyeux, M., Weinstein, A., 2010 Sep. Compulsive buying. American Journal of Drug and Alcohol Abuse 36 (5), 248–253.

Mc Elroy, S.L., Keck Jr., P.E., Phillips, K.A., 1995. Kleptomania, compulsive buying and binge eating disorder. Journal of Clinical Psychiatry 56, 14–26.

McElroy, S.L., Keck Jr., P.E., Pope Jr., H.G., Smith, J.M.R., Srakowski, S.M., 1994. Compulsive buying: a report of 20 cases. Journal of Clinical Psychiatry 55, 242–248.

Mitchell, J.E., Burgaud, M., Faber, R., Crosby, R.D., de Zwaan, M., 2006. Cognitive behavioral therapy for compulsive buying disorder. Behaviour Research and Therapy 44, 1859–1865.

Mueller, A., Mueller, U., Silbermann, A., Reinecker, H., Bleich, S., Mitchell, J.E., de Zwaan, M., 2008. A randomized, controlled trial of group cognitive-behavioral therapy for compulsive buying disorder: posttreatment and 6-month follow-up results. Journal of Clinical Psychiatry 69, 1131–1138.

Ridgway, N.M., Kukar-Kinney, M., Monroe, K.B., 2008. An expanded conceptualization and a new measure of compulsive buying. Journal of Consumer Research 35, 622–663.

CHAPTER

86

Sexual Addiction

John Bancroft

The Kinsey Institute for Research in Sex, Gender and Reproduction, Indiana University, Bloomington, IN, USA

OUTLINE

INTRODUCTION

The term sexual addiction has been widely used to describe sexual behavior that is recurrently out of control and which continues in spite of significant harmful consequences. For many years now there has been a culture of Sex Addicts Anonymous in which individuals with this problem obtain help with the same 12-step treatment programs used for alcohol- and drug-dependent individuals. There are clearly many features in common with drug addiction, and with other behavioral addictions like bulimia. But the extent to which these different patterns of addiction share the same causal mechanisms remains uncertain.

To add to the confusion, a variety of other names have been used instead of sexual addiction; for example, sexual compulsivity, impulse control disorder and, with a longer history, nymphomania applied to women and satyriasis applied to men. Hypersexuality has been used in the past and has recently been proposed as a diagnostic category for the emerging *Diagnostic and Statistical Manual of Mental Disorders* 5th edition (DSM-V). This term has always aroused controversy because of the implication that a high level or frequency of sexual behavior is necessarily problematic or pathological. Unfortunately more attention has been paid to selecting the appropriate name than to understanding the determinants of such problematic sexual behavior. The term sexual addiction will be used generically in this chapter.

In the last few years, however, some potentially key factors underlying such behavior have started to be identified, paving the way for research to establish their origins and to what extent these origins vary in importance across the range of sexual addictions. The principal emerging factors will be considered more closely in this chapter.

THE NATURE OF SEXUALLY ADDICTIVE BEHAVIOR

The most common pattern is masturbation linked to certain types of fantasy or, more often, pornography. The emergence of the Internet has added a whole new dimension. For sexually addicted men, the Internet

Principles of Addiction, Volume 1
DOI: http://dx.doi.org/10.1016/B978-0-12-398336-7.00086-3

provides a seemingly endless supply of novel pornographic material that drives their behavior, often resulting in Internet access in the workplace, as well as at home. For sexually addicted women, interaction with others on the Internet leads to cybersex. In both cases, the most likely outcome of such stimulation is masturbation. However, in some cases Internet contact leads to real-space sexual interactions. More traditional means of finding new, and typically transient, sexual partners may be used, including prostitutes and visiting strip clubs. It is important to keep in mind that with all these types of sexual expression, the majority of individuals are not out of control and do not suffer major harmful consequences as a result. In a minority, however, the urge to express their sexuality in this way may take over their lives, resulting in many hours of the addictive behavior each day, and with negative consequences for their current sexual relationship, their self-esteem, their work, and often their financial status.

THE PREVALENCE OF SEXUALLY ADDICTIVE BEHAVIOR

As yet, only one study using a representative community sample has assessed the prevalence of out-of-control sexual behaviors. In this New Zealand study of 940 men and women, all aged 32 years, participants were asked about sexual fantasies, urges, or behavior that they regarded as out of control during the previous year. Nearly 13% of men and 7% of women reported such experiences, although only 3.8% of men and 1.7% of women believed that these experiences had interfered with their lives. With these men, paying for heterosexual sex and same-sex attraction and behavior were typical themes. With the women, a high number of male partners, concurrent sexual relationships, sex with a partner met on the Internet and same-sex attraction and behavior were typical. It is not possible from this study to assess the duration of these problematic patterns, and in many cases they may well have been transient.

THE NEUROPSYCHOLOGY OF SEXUAL AROUSAL AND DESIRE

In much of the writing on sexual addiction, and particularly that on so-called hypersexuality, a high level of sexual desire is considered the primary problem. However, the individual acting out his or her sexual addiction is likely to be in a sexually aroused state. The relation between sexual arousal and sexual desire therefore needs to be considered. Sexual arousal is a complex state that involves (1) information processing, both automatic or unconscious, and attentional or conscious; (2) incentive motivation; (3) general arousal; and (4) genital response. Sexual desire involves some or all of these components to some extent, although often this may be restricted principally to incentive motivation plus relevant information processing that identifies the target of the incentive.

The incentive motivation component is probably similar to that for other types of appetitive patterns, and is the aspect of sexuality most directly relevant to the basic mechanisms of drug addiction. However, there is as yet no clear evidence of an aberrant function of this component in sexual addiction of the kind that has been demonstrated with the chronic use of drugs of addiction.

When we look more closely at other components of sexual arousal and desire some important mechanisms become evident, which serve to distinguish the sexual experience as having certain relatively unique characteristics. This is most apparent in relation to genital response, the most fundamentally sexual component of sexual arousal, such as penile erection in the male, and clitoral and vulval tumescence in the female. It is important to acknowledge that specialized inhibitory mechanisms are involved.

Inhibitory Mechanisms

The occurrence of genital response and, with it, the characteristic sequence of general arousal, incentive motivation, and relevant information processing, which combine to produce the experience of sexual arousal, are determined by the interaction between excitatory and inhibitory mechanisms. The Dual Control Model was developed at the Kinsey Institute as a way of conceptualizing this complex process. This model postulates that in the presence of sexual stimuli, either external (e.g. pornography) or internal (i.e. sexual fantasy or thoughts), excitatory mechanisms are activated, but are balanced by inhibitory mechanisms. Whether sexual arousal or sexual desire occurs on any particular occasion, depends on this interaction. At least three types of sexual inhibition have been identified: (1) reactive inhibition that results from information processing that identifies negative implications of the stimulus; (2) inhibitory tone, which normally prevents genital response and other more specifically sexual components of sexual arousal from occurring, and which needs to be switched off if sexual arousal is to occur; and (3) the postorgasmic refractory state, which is much more evident in men than in women, and which combines a temporary reduction of excitation potential with high levels of inhibition. Each of these inhibitory mechanisms is based on automatic (unconscious) information processing. A conscious process of self-regulation, which is not restricted to

sexual situations, but which is relevant to how we manage our sexual lives, will be considered later in this chapter.

Impaired Inhibition of Sexual Arousal

A fundamental assumption of the Dual Control Model is that individuals vary in their propensity for both sexual excitation and sexual inhibition. Psychometrically validated measures of these two propensities have recently been established, and across a range of large samples studied, close to normal distributions of sexual excitation scores and sexual inhibition scores have been demonstrated. Men, on average, score significantly higher on sexual excitation than women, and women, on average, score significantly higher on sexual inhibition than men, although with considerable overlap of the male and female distributions.

This normally distributed individual variability in both men and women has been taken to mean that for most individuals their propensity for sexual excitation and sexual inhibition are appropriate and adaptive, with normal inhibitory mechanisms helping to keep them out of trouble in their sexual lives. In contrast, the more extreme ends of these distributions are likely to be problematic; those with low propensity for sexual excitation and/or high propensity for sexual inhibition are likely to experience difficulties with their sexual responses, or sexual dysfunctions; those with high propensity for sexual excitation and/or low propensity for sexual inhibition are likely to have problems with inappropriate or out-of-control sexual behavior, such as sexual addictions or sexual risk taking. So far, this recently developed approach has been used to explore determinants of high-risk sexual behavior, and the predictions have been well supported in both heterosexual and gay men and in heterosexual women. More limited evidence has shown that men and women who have sought treatment for sexual addictions have a high propensity for sexual excitation, with the men also having a low propensity for sexual inhibition. Furthermore, in a small Kinsey Institute study of self-defined sex addicts, we found sexual inhibition to be more relevant to addictive behaviors involving other people than to masturbatory patterns. This difference needs to be explored further.

The lack of support for low sexual inhibition in women with sexual addictions, reported above, should not as yet be considered conclusive as the measures used were designed initially for men, and may not have asked the best questions about low sexual inhibition in women. The case has been made that inhibitory mechanisms are more crucial and hence better developed in women. A new measure, called the Sexual Excitation/Sexual Inhibition Inventory for Women (SESII-W), has been developed, but has not, as yet, been used in relation to sexual addictions in women.

The Relevance of Negative Mood

It has been widely asserted in the literature that negative mood, both depression and anxiety, is relevant to sexual addiction. This has been reported as comorbidity, with sex addicts being more likely to have a history of mood disorder, but also as a direct relationship, with the addictive behavior functioning to improve negative mood.

This, at first sight, seems surprising as, for most people, negative mood in terms of both depression and anxiety is associated with reduced sexual interest and/or response. However, a consistent finding across clinical studies of mood disorders has been that a minority of patients report a paradoxical tendency for sexual feelings to be increased during negative mood states. As yet, little attempt has been made to explain these unusual patterns.

Recently, research at the Kinsey Institute has focused on this paradoxical pattern in nonclinical groups. A simple instrument, the Mood and Sexuality Questionnaire (MSQ), assessed sexual interest and sexual response during states of both depression and anxiety, with four questions: "When you have felt depressed (or anxious/stressed) what typically happens to your sexual interest (or response)?" In studies of heterosexual men and gay men, and heterosexual women, of those who had experience of negative mood states, around 10% of men and women reported increased sexual interest when depressed, and around 20% when anxious or stressed.

Are such paradoxical mood–sexuality relationships relevant to sexual addictions? In the small Kinsey Institute study of 31 sex addicts, 27 participants (87%) stated that their sexual acting out was predictably affected by their mood; 17 of them reported being more likely to sexually act out when depressed, and 19 when anxious or stressed. Eleven subjects reported this in states of both depression and anxiety. Only two men were less likely to act out when depressed, and no one reported this tendency in relation to anxiety. As a group, their scores on the MSQ were significantly higher than those of the controls ($P < 0.001$), indicating their greater likelihood of being sexually responsive in negative mood states, and they also scored significantly higher on a trait measure of proneness to depression. An association between negative mood and sexually addictive behavior therefore appears to be highly relevant.

Qualitative data from a nonclinical study indicate a more complex relationship between depression and sexuality, and a comparatively simple relationship

with anxiety, most commonly expressed by masturbation. Two mediating patterns are of particular interest:

1. In states of depression one might pursue sexual contact with another person to satisfy depression-related emotional needs, such as making personal contact through sex, feeling validated by another person, and enhancing one's self-esteem by feeling desired by another person. These are direct examples of sex used as a mood regulator.
2. The tendency for sexual interest and sexual arousal to be increased in negative mood states characterized by increased arousal (such as anxiety or stress) may result from excitation transfer.

Both of these patterns are easier to understand if we return to the Dual Control Model. Our original studies of mood and sexuality found that in heterosexual men, the measures of propensity for sexual inhibition were negatively predictive of the paradoxical relationship between depression or anxiety and sexuality, with the measure of sexual excitation propensity being positively but less strongly predictive. In women, only the paradoxical relationship with anxiety was predicted and only by the measure of propensity for sexual excitation. The explanation for this apparent gender difference is not yet clear. However, we can reasonably assume that in men, a normal propensity for sexual inhibition would preclude a paradoxical sexual response while experiencing negative mood. It is therefore evident in the individual man with a low propensity for sexual inhibition that the paradoxical mood/sexuality pattern would allow the types of mood regulation described above.

The second mediating pattern involves excitation transfer associated with states of anxiety or stress. The concept of excitation transfer infers that arousal induced by one type of stimulus can be transferred to enhance the arousal induced by another. Thus if an individual is aroused because they are feeling anxious, this arousal may get transferred to augment arousal in response to a sexual stimulus. However, this would only occur if there was less than normal inhibition of sexual arousal in states of anxiety. But this may be the case in persons prone to sexually addictive behavior. Transfer of anxiety or stress into sexual arousal creates a strong incentive to pursue sexual release through orgasm, leading to the establishment of learned or even conditioned associations between negative mood and sexual arousal. However, the subsequent recognition of this as a recurring and out-of-control pattern induces further negative mood. This pattern is most likely to be manifested in solitary or masturbatory patterns of behavior.

If this paradoxical mood/sexuality relationship is of relevance to some types of sexual addiction, this raises the question of why some individuals have the capacity for these atypical and potentially problematic interactions between mood and sexuality but the majority do not. A negative correlation between MSQ score and age has been found in heterosexual men, which probably means that such paradoxical patterns are more common in younger men and lessen with age. This raises the question of when such an association becomes established? It is a distinct possibility that this occurs during childhood or early adolescence as a consequence of early experiences, such as child sexual abuse or induced guilt about masturbation, that combine sexual response with negative mood. This hypothesis would not be difficult to test with retrospective research.

It has also been postulated that the use of sexual behaviors such as masturbation to alleviate emotional pain may reflect an intimacy dysfunction resulting from child sexual abuse or neglect. An early establishment of a paradoxical mood/sexuality pattern could interfere with normal sexual development, creating a barrier to the incorporation of one's sexuality into close, intimate sexual relationships. It is therefore noteworthy that in a study of heterosexual men, those in exclusive monogamous relationships had lower MSQ scores; that is, were less likely to report this paradoxical pattern.

As yet, we do not have an adequate explanation for these paradoxical mood/sexuality relationships, and the use of the Dual Control Model only provides partial explanation. Further research is needed on this important issue.

SEXUAL ADDICTION AS AN OBSESSIVE-COMPULSIVE DISORDER

The term compulsive has been widely used as a descriptor of out-of-control sexual behavior, without reference to whether such compulsivity relates to obsessive-compulsive personality disorder. In DSM-IV, out-of-control sexual behavior is excluded from the obsessive-compulsive (OCD) category on the grounds that "the person usually derives pleasure from the activity and may wish to resist it only because of its deleterious consequences." Compulsive OCD thoughts often do have sexual content, but are typically accompanied by negative mood and no sexual arousal. Most people with obsessive-compulsive disorder combined with a propensity for mood disorders are likely to experience a decline in sexual arousability during negative mood states, as is the norm. But exceptions do occur occasionally. An example, which has been reported in the literature, is when obsessive-compulsive symptoms include intrusive sexual thoughts accompanied by penile erection. The awareness of the erection intensifies the

anxiety, thus reinforcing the process. This could occur in an individual with obsessive-compulsive tendencies and also a capacity for a paradoxical mood/sexuality relationship, resulting in an atypical, sexualized type of compulsive pattern. If so, one would expect to find evidence of other obsessive-compulsive phenomena in such individuals.

Obsessive-compulsive urges are characterized by their egodystonic nature, typically associated with an attempt to resist them. In the Kinsey Institute study of sex addicts, referred to earlier, 15 subjects were asked whether they tried to resist the urge to act out or whether at the time it was something they genuinely wanted to do. Eleven men and one woman said that they tried to resist, but most did not give a convincing description of resistance (e.g. "I tell myself not to do it, but I do it anyway"). The two most convincing accounts of resistance were from men with obsessive-compulsive disorder. In both cases the sexual acting out was masturbation. One man had intrusive thoughts about teenage boys, or a compulsion to look at pictures of them. This was associated with considerable guilt and resistance. He masturbated to escape from these thoughts and to achieve a very transient calming effect. However, this was followed quickly by renewed guilt and depression. His resistance was to the intrusive thoughts about boys rather than the masturbation. The other man described ruminative preoccupation with sexual thoughts, which led to masturbation followed by the need to shower because of the "dirtiness" of the act.

A few studies have looked for evidence of obsessive-compulsive disorder among sex addicts, usually finding around 15% in this category.

Thus, some forms of out-of-control sexual behavior can be appropriately regarded as atypical obsessive-compulsive phenomena, but this clearly applies to a small minority of those regarded as sex addicts.

A Female Problem

Of interest, is Persistent genital arousal disorder (PGAD) a uniquely female problem that has only recently received attention in the literature. Women with this problem report their genital response and arousal as continuous, overwhelming, and distressing, in spite of repeated masturbation in an attempt to terminate it. In an Internet survey, 76 women with PGAD were likely to be depressed, to report panic attacks, and to monitor their physical sensations in an obsessive-compulsive manner. Two gender differences in sexual response may help us explain this peculiarly female phenomenon. First, women have much less refractory inhibition after orgasm than men; thus, orgasms are less likely to have a limiting effect. Second, a uniquely female aspect of genital response is the automatic increase in vaginal blood flow in the presence of any sexually relevant stimulus, whether or not the woman finds the stimulus appealing. Whereas this vaginal response is often interpreted as a manifestation of sexual arousal in women, it is not predictably associated with subjective arousal, and there is no parallel to this response in the male. However, a woman may be aware of it, and this awareness, in certain women, may initiate a process of monitoring, leading to the augmentation and persistence of the response because of the impact of negative emotions. As yet, the relationship of PGAD to obsessive-compulsive personality disorder is uncertain, but it can clearly lead to out-of-control sexual behavior, particularly masturbation, in women.

FAILURES OF SELF-REGULATION

Apart from the more physiological mechanisms that have been considered as influencing sexual activity, we should also consider the more cognitive approach to self-regulation, which is not, of course, restricted to self-regulation of sexual behavior, and about which there is a substantial literature.

Three components of self-regulation have been described: (1) standards, (2) monitoring, and (3) the operative phase of regulation. Standards are of interest; in particular, the capacity for strict standards to undermine self-regulation. In the Kinsey Institute study of sex addicts, religion was very important for five men. For such individuals, the unquestionable moral unacceptability of most types of sexual behavior conflicts with their sexual impulses, undermining any sensible pattern of regulated sexual behavior. For example, for an individual who believes masturbation to be evil, and who has strong impulses to masturbate in response to images on the Internet, a regulated pattern of masturbation is not an acceptable option, although for most people this can be a responsible way of dealing with their sexual needs. Highly restrictive sexual attitudes can result in inability to conform, starting off a cycle of guilt, mental pain, and the urge to act out.

Self-monitoring is fundamental to effective self-regulation. Alcohol use, as well as fatigue and stress, impairs normal monitoring. This has been called alcohol myopia and euphoric recall and occurs when intoxication causes attention to focus on the positive, sexually arousing, or rewarding aspects of the situation and away from negative consequences and the associated inhibition of arousal. Sexual arousal can be likened to alcohol intoxication in this respect.

Another way to conceptualize this aspect of monitoring is as transcendence or focusing one's awareness

beyond the immediate situation, so that distal concerns or consequences are kept in mind. In the Kinsey Institute sex addicts study, we unexpectedly found indications of a dissociative tendency that could contribute to out-of-control behavior by undermining, if not eliminating, this transcendence. When asked to describe a typical state of mind while acting out, 14 men (45%) gave descriptions suggestive of some degree of dissociation. The following are illustrative examples, each from a different subject: "… an overpowering drive … nothing else is under consideration"; "… numb, completely zoning out, not present, not conscious of reality"; "… trance-like state … kills time and pain …. numb like a dream"; "… feel detached from what is happening"; "… like a drug to numb out." It is possible that such explanations may be post hoc justifications for the behavior, but dissociation has not been explored in the relevant literature and warrants closer study. Do people with out-of-control sexual behavior show more dissociative tendencies in general?

Impairment of self-regulation is clearly relevant to out-of-control sexual behavior, at least in some cases.

TREATMENT

Given the uncertainty and likely variability of the etiology of sexual addictions, it is not surprising that there are no clear treatment guidelines. When the sexual addiction is related to negative mood, a cognitive-behavioral approach may be helpful. The patient is encouraged to look for a typical sequence of events that leads from awareness of the negative mood state to the out-of-control sexual behavior. This may allow identification of stages in the sequence when some alternative and less problematic mood-regulating behavior can be substituted.

An alternative approach is pharmacological. Most attention has been paid to selective serotonin reuptake inhibitors (SSRIs), which have the potential for helping in two ways: by improving negative mood and by inhibiting sexual response. A more radical approach is hormonal, using anti-androgenic drugs such as cyproterone acetate or medroxyprogesterone acetate, which suppress sexual appetite and have been widely used in the management of sexual offenders. However, these may have negative effects on mood. In general, it is preferable to integrate any pharmacological methods into a cognitive-behavioral approach.

CONCLUSIONS

The limited evidence so far available points to a variety of mediating mechanisms that lead to uncontrolled sexual behavior. It remains a possibility that some dysregulation of the incentive motivation system occurs in some cases, comparable to that found in drug addiction. But until we have evidence of this, it is probably best to consider the term sexual addiction as an analogy. Impairment of the inhibitory mechanisms, a paradoxical mood/sexuality relationship, and deficiencies in the self-regulation mechanisms, plus the impact of obsessive/compulsive disorder are all of potential relevance, with the likelihood that when these are better understood we will arrive at a number of subtypes with different etiologies and possibly different methods of treatment. The concept of hypersexuality, currently prominent in the DSM-V recommendations, should be seen as a contributory factor (i.e. high levels of sexual drive or arousability) rather than a sufficient explanation. The common theme, however, is lack of control. In these circumstances, it may be best to use the umbrella term out-of-control sexual behavior to cover this range of problems until we have an advance in our knowledge. The good news is that we seem to be at a point when useful research can be carried out to test the various possibilities addressed in this chapter.

SEE ALSO

Internet Addiction: Cybersex, Impulsivity, Disinhibition, and Risk Taking in Addiction, Alcohol and Sexual Violence, Alcohol, Sexual Risk Taking, and Sexually Transmitted Infections, Alcohol's Effects on Sexual Arousal and Sexual Functioning

Glossary

Dissociation when normally connected mental functions become disconnected.
MSQ mood and sexuality questionnaire.
PGAD persistent genital arousal disorder.
SSRIs selective serotonin reuptake inhibitors, a category of antidepressant medication, e.g. fluoxetine, paroxetine.
Anti-androgen a substance that blocks or counteracts the hormonal effects of androgens such as testosterone.

Further Reading

Bancroft, J., Graham, C.A., Janssen, E., Sanders, S.A., 2009. The dual control model: current status and future directions. Journal of Sex Research 46, 121–142.
Bancroft, J., Janssen, E., Strong, D., et al., 2003. The relation between mood and sexuality in heterosexual men. Archives of Sexual Behavior 32, 217–230.
Bancroft, J., Vukadinovic, Z., 2004. Sexual addiction, sexual compulsivity, sexual impulsivity or what? Towards a theoretical model. Journal of Sex Research 41, 225–234.

Baumeister, R.F., Heatherton, T.F., 1996. Self-regulation failure: an overview. Psychological Inquiry 7, 1–15.

Goodman, A., 1997. Sexual addiction. In: Lowinson, J.H., Ruiz, P., Millman, R.B., Langrod, J.G. (Eds.), Substance Abuse: A Comprehensive Textbook. Williams & Wilkins, Philadelphia, pp. 340–354.

Kafka, M.P., 2010. Hypersexual disorder: a proposed diagnosis for DSM V. Archives of Sexual Behavior 39, 377–400.

Leiblum, S., Seehuus, M., Brown, C., 2007. Persistent genital arousal: disordered or normative aspect of female sexual response. Journal of Sexual Medicine 4, 680–689.

Skegg, K., Nada-Raja, S., Dickson, N., Paul, C., 2010. Perceived "out of control" sexual behavior in a cohort of young adults from the Dunedin multidisciplinary health and development study. Archives of Sexual Behavior 39, 968–978.

Winters, J., Christoff, K., Gorzalka, B.B., 2010. Dysregulated sexuality and high sexual desire: distinct constructs? Archives of Sexual Behavior 39, 1029–1043.

Zillman, D., 1983. Transfer of excitation in emotional behavior. In: Cacioppo, J.T., Petty, R.E. (Eds.), Social Psychophysiology: A Sourcebook. Guilford Press, New York, pp. 215–240.

CHAPTER

87

Areca Nut, Betel Quids, and Associated Products

Adam R. Winstock

South London and Maudsley NHS Trust, Kings College London, London, UK

OUTLINE

INTRODUCTION

The areca nut, often incorrectly referred to as betel nut (betel refers to the leaf that the nut is often wrapped in), and its products are the fourth most commonly used psychoactive substance in the world after tobacco, alcohol, and caffeine. Like all such substances its use traverses age, class, and culture, with an estimated 600 million users globally. Areca is a masticatory and is chewed often in combination with tobacco.

The areca nut is the seed of the fruit of the oriental palm Areca catechu and is the core constituent of a variety of raw and purified products (see Fig. 87.1). The nut is a hard brown oval kernel, about the size of a plum stone. The areca palm is not only cultivated for the nut but also for its husk, which is used in paper and insulating wool production. The nut may be eaten raw or refined in a number of ways including boiling in water, roasting, or curing.

It is used most commonly within the Indian subcontinent (India, Nepal, Pakistan, and Bangladesh), but is also prevalent in Taiwan and Southeast Asia (Indonesia, Thailand, Philippines, Cambodia, Loas, Guam, Malaysia among others), the Pacific rim and southern China. Areca is known by different names in different countries. For example, in India, it is commonly referred to as supari, in Papua as daka, and in Thailand as gua. Migration and displacement have carried its use more widely in recent decades and consumption is now seen in parts of Africa, Australasia, Europe, and Northern America.

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00087-5

FIGURE 87.1 Raw dried nut. *Copyright permission applied for www.tootoo.com.*

Although areca can be characterized as a psychomotor stimulant its effects are quite subtle, more comparable to those of coffee or chewed coca or khat leaves than amphetamine. Anecdotally the nut also has significant medicinal properties ranging from anti-helminthic (in Ayurvedic medicine) to astringent, dentifrice, aphrodisiac, and digestive enhancer. The major risk associated with the use of areca nut is the dose dependent risk for the development of premalignant oral lesions including leukoplakia, submucous fibrosis, and oral squamous cell carcinoma.

For such a widely used substance, it is perhaps surprising as to the paucity of literature within the addictions field regarding its use. Possible explanations include its wide cultural acceptability, relatively mild psychoactive effects, uncertainty over its abuse potential, and the absence of overt public health consequences beyond those to the individual consumer.

HISTORY

The first reports of areca nut use date back to the fifth century BC where legend tells of a princess from the Pali region of India giving a gift of areca to her nurse. Areca is referred to in ancient Sanskrit scriptures and its medicinal properties were first described by the scholar Vagbatta in 500 AD. Throughout diverse cultures the nut has a rich history with the origins of its use embedded in religious, ceremonial, medical, and social activities. Such functions are supported by the discovery of artifacts including elaborate utensils for its preparation and serving, which have been identified throughout many ancient civilizations. Anthropological texts also describe a wide range of functions from improving stamina and work performance to supporting fertility.

EPIDEMIOLOGY

The largest areca using population in the world is India, with a prevalence of 20–40% in some regions, followed by Indonesia, China, and Taiwan. In many counties children are first introduced to areca by family members in social or ceremonial gatherings. In most cultures use is more prevalent among men, though the gender ratio of use varies widely between cultures. Epidemiological studies of use in south Asia and India generally find that use is more common among older people though in some countries such as Micronesia the age of onset is much earlier (average 12 years) resulting in 10% having oral pathologies by their teens.

Use tends to be higher in those with lower levels of education and those engaged in manual forms of employment. Other risk factors for use include prior smoking of tobacco, parental use of tobacco or areca chewing, living in rural and remote areas, peer pressure, social and cultural identification and aggressive marketing and promotion to risk groups such as truck drivers in Taiwan and school children in India. Recent genetic studies have also identified specific alleles for mono amine oxidase-A (MOA-A) activity that may confer elevate risk for heavy betel quid (BQ) use. The level of use however varies widely between cultures, preparations, and type of user. Most studies of regular users report typical daily users chewing 1–10 nuts day^{-1} in 4–10 sessions. Heavy users may chew more than 40 nuts per chewing continually throughout the day, even sleeping with a quid in their mouth at night.

Some countries such as Thailand, Malaysia, Singapore, and Cambodia have seen a fall in recent decades in the prevalence of areca nut. In Thailand this has been reflected by a decline in oral cancer rates. Over the same period however Thailand has seen a corresponding increase in the use of alcohol and amphetamines. Conversely Taiwan has seen a recent increase in the prevalence of areca use. This appears related to the freeing up of the tobacco market in Taiwan in the 1908s. Studies indicate that in Taiwan tobacco smoking acted as a gateway to chewing, with most chewers in Taiwan also being tobacco smokers. The escalation of both habits (not helped by the frequent location of shops selling betel, tobacco, and alcohol) has led to a fourfold increase in the rate of oral cancer in the past two decades.

The effects of migration on use are inconsistent. Migrant communities may record an increase in use where continued use assists in defining ones cultural heritage. Equally among second-generation immigrants there may be a transition to more local patterns of substance use that are often associated with greater risk of harm.

PREPARATIONS (SEE TABLE 87.1)

Areca nut is traditionally used in combination with the leaf of the piper betel vine (a member of the piperaceae or pepper family). The leaf contains aromatic

TABLE 87.1 Major Preparations, Composition, and Region

Preparation	Composition	Region
Areca nut	Unripe, ripe, baked, roasted, boiled, fermented	Various
Betel quid (pan)	Areca (ripe), betel leaf, slaked lime, flavorings +/− tobacco	India, Taiwan
Lao-hwa quid	Split unripe nut sandwiched between flower of Piper betel Linn, red lime	Taiwan
Stem quid	Split unripe nut sandwiched between flower of Piper betel Linn, white lime	Taiwan (aborigines)
Mainpuri	Tobacco, slaked lime, camphor and cloves	India
Supari	Roasted flavored nut pieces	India−Uttar Pradesh
Mawa	Areca, slaked lime tobacco	Northern India
Gutkha	Areca nut, lime, catechu, flavorings, fragrance, and tobacco	Gujarat−India and immigrant populations, for example, in the UK and USA
Pan masala	Areca nut, lime, catechu, flavorings, and fragrance	

FIGURE 87.2 (a) Typical composition of betel quid: Betel leaves (A) with slaked lime (B), areca nut (C) and tobacco (D). (Copyright applied for International agency for the research on cancer (http://screening.iarc.fr) WHO) (b)Alternative – Constituents of a betel quid. *Copyright permission applied for – www.thedevilslastsupper.com*

phenols including chavicol, allylpyrocathechol, chavibetol, and cadinene. The most common and traditional preparation of areca nut is in the form of the "betel quid" (see Fig. 87.2(a) and (b)), comprising the betel leaf, sliced areca nut, and slaked lime (aqueous calcium hydroxide). Slaked lime (from burnt coral, shell fish or lime stone) increases the pH of saliva facilitating the extraction of arecoline and guvacoline, which are then absorbed transbuccally and a to lesser extent sublingually and orally into the bloodstream. The addition of lime to the quid also increases the generation of mutagenic oxygen-free radicals and is itself caustic to oral mucosa, rendering it more susceptible to carcinogens. Quids may be prepared commercially at small stalls or be prepared by the user themselves, the individual taking all that is needed with them in a pouch or tin.

Beyond this basic recipe however, there are myriad variations in the treatment of nut (it may be used ripe or unripe, raw, baked, boiled, roasted, or fermented all of which reduce the tannin content, improving palatability), its additives. These include catechu (an astringent, reddish-brown substance that is often smeared on the betel); kratom (which has been used as an opiate substitute in Thailand and Malaysia); flavoring ingredients (which can include spices such as cardamom, cloves and cinnamon, and various condiments); various types of lime (red or white). Variation is also seen in the precise preparation of the quid (see Fig. 87.3). Uniquely in Taiwan in addition to the leaf, the safrole containing betel inflorescence is used (though usually not with tobacco).

Perhaps the most significant variation in composition from a health and abuse potential perspective is the common addition of tobacco to the quid or other areca preparations such as Mawa (areca, tobacco, and slaked lime), which is seen in the Gujarat province of India. The tobacco itself may be dried, fermented, or powdered. It is probable that tobacco became available to chewers through its introduction into India by Europeans in the sixteenth century. Cultural variations are also seen in the administration of the various products. For example, quid chewers in Papua New Guinea apply the lime separately to the corner of the mouth. In parts of India the nut has more recently been commercially prepared as Supari – roasted and sweetened small pieces of nut served as a digestive.

The use and formulation of areca like other naturally occurring psychoactive substances have been impacted upon changes in technology and lifestyle. Traditional preparations of the nut were perishable and required a fresh source of the nut, limiting it appeal in

FIGURE 87.3 Woman preparing betel quids. *Copyright Adam R Winstock*

modernizing cultures. Advances in purification, packaging, and branding have led to the proliferation of a wide range of dry, relatively nonperishable, industrially manufactured, freeze dried formulations of the nut. India has led the way in the development and marketing of these products, which are sold in heavily branded foil sachets or tins (see Fig. 87.4). Typically preparations comprise a combination of powdered areca nut, lime, catechu, flavorings, and fragrance. Without tobacco the product is known as pan masala. When the composition includes tobacco the product is referred to as gutka. The addition of sweeteners fragrances and inconsistent product labeling mean tobacco-containing products represent a covert and unregulated entry point for young people to nicotine dependence. Although banned in many states in India and unavailable in much of Southeast Asia, these products are exported globally and are still widely available.

Distribution and sale of the various preparations also vary widely between cultures from road side sellers and corner shops, to the scantily clad Bin Lang ("betel quid barbies") girls in Taiwan, to online sales supporting global marketing. Of particular concern is the aggressive marketing of prepackaged tobacco-containing products targeting younger users. The commercial value of areca products worldwide is likely to be worth hundreds of millions of dollars each year.

FIGURE 87.4 Lao-hwa quid. *Copyright Adam R Winstock 2005*

ROUTE OF ADMINISTRATION

Areca is a masticatory (chewed substance). The process of chewing breaks down the nut and other organic material releases arecoline and other active compounds into the saliva. Typically after a few cycles of mastication the folded leaf package ("quid") is then placed in the mouth and held against the oral mucosa of the buccal cheek and molar teeth. With the initial taste being rather bitter, many users spit out the product of the first masticatory cycle. Periodic rechewing supports further extraction of the juice. Although chewing leads to some gastrointestinal absorption (subject to first-pass metabolism), the primary route of absorption is likely to be direct systemic absorption through the buccal and sublingual mucosa, leading to an onset of effect in the majority of chewers within 5 min, with a duration of action of 2–3 h. Typical red spittle patterns frequently marble the streets of many cities where the habit is common, often accompanied by "no spitting" signs.

SOUGHT AFTER EFFECTS (SEE TABLE 87.2)

Areca shares much in common with nicotine, inducing a relatively mild but reinforcing mix of

TABLE 87.2 Sought After and Physiological Effects

Sought after effects	Physiological effects
Increase in mood and well-being	Salivation
Reduction in hunger	Sweating
Improvements in concentration/performance	Offset fatigue
Improved socialization	Nausea
Hot bodily sensation (a feeling of warmth)	Increased heart rate
Pain relief	Pupillary constriction
Oral hygiene astringent and breath fresher	Broncho-constriction
Facility bowel movements and digestion	Vasodilatation and facial flushing
Calming/relaxant	Possible anti-helminthic

stimulant and anxiolytic effects. Its subtle stimulant effects with a gradual onset of action related both to the low potency of its active ingredients and route of administration – mean that it has more in common with coffee or coca leaves than with cocaine or amphetamine. As with other stimulant drugs the precise effects will be dependent upon the dose consumed, whether or not tobacco is also used and the pharmacogenetic profile and tolerance of individual. Consistent with its history of ceremonial use and distribution at social gathering, users report a mild elevation in their sense of well-being and an increase in sociability.

PHARMACOLOGY AND PHYSIOLOGY

The nut contains a number of psychoactive alkaloids, of which arecoline (methyl-1,2,5,6-tetrahydro-1-methyl-nicotinate), a volatile oil that resembles nicotine in both effect and structure, is the one present in the greatest quantity. Arecoline is probably responsible for most of the sought-after effects of areca nut chewing but is also responsible for its cytotoxic and genotoxic properties. The nut also contains a number of other potentially psychoactive compounds including the alkaloids arecaidine, guvacoline, guvacine, and a number of polyphenolic compounds and tannins.

The nut has complex physiological effects and activates both the sympathetic and parasympathetic nervous systems. Chewing leads to a transient increase in sympathetic activity (with elevations in plasma noradrenaline and adrenaline levels), manifested as an increase in stamina, heart rate, skin temperature, and a reduction in hunger and choice reaction time. Electroencephaolgraphic are also consistent with increased sympathetic activity studies with increases in both alpha and particularly beta rhythms.

Arecoline is also a potent cholinergic agonist and leads parasympathetic activation resulting in bradycardia, pupil constriction (overall there are insignificant effects on pupil size because of parallel sympathetic activity), and increases in salivation and smooth muscles stimulation consistent with subjective pro-digestive properties. Activity at muscarinic and nicotinic receptors may underlie the subjective and now laboratory confirmed reports of improved learning in healthy volunteers. The use of arecoline has also been investigated as a potential cognitive enhancer in those with dementia. Specific areca alkaloids also act as competitive inhibitors of γ-aminobutyric acid receptors in the brain, cardiovascular system, and pancreas. Areca is also rich in copper, which has been implicated in the pathogenesis of oral submucous fibrosis (OSF).

SHORT-TERM ADVERSE (SEE TABLE 87.3)

A combination of its cholinergic activation and psycho stimulant effects mean that areca may occasionally result in adverse gastrointestinal, cardiovascular, and neurological experiences. These are generally short-lived, self-limiting, and will be more common in naïve users where tolerance has not been developed (see Table 87.3)

Occasionally harms may arise from contamination of the nut or other quid constituents by naturally occurring harmful substances, for example, aflatoxin (a naturally occurring mycotoxin that is produced by many species of the fungus Aspergillus).

LONG-TERM EFFECTS (SEE TABLE 87.4)

Excessive long-term use of areca nut and BQ has been associated with a number of health-related harm issues most seriously oral cancers. The risk of long-term harm is related primarily to frequency and duration of use as the concurrent use of tobacco. The use of tobacco and areca results in synergistic harms which are dose dependent.

OSF is a high-risk premalignant condition characterized by the insidious development of fibrous white bands representing blanching and stiffening of the oral mucosa and oropharynx (see Fig. 87.5). Symptoms include oral pain and hypersensitivity, characterized by a burning sensation in the mouth, trismus and reduction in mouth opening, which can be severe, limiting speech, and mastication. The disease is nonreversible and malignant transformation is seen in 3–8%. The risk of OSF is in part related to the preparation used. For example, unripe areca nuts have higher tannin levels than ripe nuts and are more likely to result in OSF. More worryingly, freeze dried products such as pan masala, guthka, and mawa (areca and lime) that have high concentrates of areca nut per chew, appear to cause OSF more rapidly (in some cases after just 3–4 years of chewing) than conventional raw BQ preparations that contain smaller amounts of areca nut.

TABLE 87.3 Short-Term Adverse Effects

Tachycardia/palpitations hypotension and bradycardia
Tremor
Sweating
Dizziness
Diarrheas, nausea and vomiting
Shortness of breath, worsening of asthma

TABLE 87.4 – Long-Term Adverse Effects of Areca Chewing

Oral	Discoloration of teeth and gums Mouth ulcers, mucositis, gingival disease, gum recession, periodontitis, molar wearing and cavitation Leukoplakia Eyrthoplakia Submucous fibrosis 10-fold increase in risk of oral squamous cell carcinoma
CVS	Inconsistent data supports increased risk of cardiovascular disease and cerebrovascular accidents
Metabolic	Hyperglycaemia and increased risk of type 2 diabetes and central obesity
Gastrointestinal	Peptic ulceration Liver cirrhosis and hepatocellular carcinoma
Pregnancy outcomes	Low birth weight
Psychiatric	Dependence

Leukoplakia is an asymptomatic condition characterized by patches of keratosis (over growth of keratin) visible as white patches inside the mouth. It is seen in about 10–20% of chewers. Although areca alone can cause leukoplakia, the risk is elevated in those who also chew tobacco.

Oral cancer. Areca nut has been classified as a group 3 human carcinogen. Its potential for causing malignancy is dose dependent. Among high prevalence chewing populations more than 50% of oral cancers are attributable to areca use. There appears to be a synergistic dose dependent interaction between areca and tobacco with an elevated risk of oral pathologies and cancers (typically oral squamous cell carcinoma). A three- to tenfold increase in cancer risk is seen among smokers who are also areca chewers. Certain genetic polymorphisms are also associated with an increased risk of developing areca-related malignancy. As such inquiry about a family history of chewing and oral cancer is recommended. Areca is also a risk factor for other head and neck cancers as well as hepatocellular carcinoma.

FIGURE 87.5 Pan masala and guthka.

OTHER HEALTH CONSEQUENCES

More recently research has identified areca use as a risk factor for a range of metabolic disorders including type 2 diabetes. There are no reports linking areca nut use and any specific psychiatric disorders, though it should be stated that there has been little search in this area. There are no reports of acute psychosis developing following the use of areca. In small but a representative population study, among patients with schizophrenia areca use was associated with milder symptomatology and no increase in extra pyramidal symptoms.

EFFECTS ON PREGNANCY AND THE NEWBORN

Arecoline crosses the placental barrier. Adverse pregnancy outcome has been reported from a number of studies. Controlling for tobacco and alcohol use a large Taiwanese cohort study found that mothers who chewed areca through pregnancy have poorer birth outcomes, including lower birth weight, reduced birth length, and a lower ratio of male to female offspring. Another study from Papua New Guinea has confirmed the association of areca nut usage and low birth weight. Of interest this later study reported that one-third of women indicated that a reason they had continued to chew through pregnancy was to prevent morning sickness. In general babies appear healthy in most perinatal respects. There has been single case report of suspected neonatal withdrawal.

ABUSE LIABILITY AND DEPENDENCE

The areca nut demonstrates a number of characteristics common among substance of abuse. Preparation and use is often ritualized, effects are reinforcing, and chronic habitual use is common across cultures. About 50% of users are daily chewers and dependence has been reported to be seen in 15–60% of users. One of the first references to its addictive potential was in a research paper entitled "Betel mania" exploring use among Cambodian refugees in the United States in the early 1990s.

Rates of dependence are higher among those chewing refined products and those mixing areca with tobacco. Chewers also smoke more than nonchewers. Tolerance

is seen, with users increasing their dose. This is supported by laboratory studies demonstrating that habitual chewers show a less marked increase in pulse than occasional chewers and reports of naïve chewers being more likely to report adverse symptoms such as tremor, dizziness, gastrointestinal upset, and nausea. Although chewing delivers a more gradual onset of effect than smoking, experience from smokeless tobacco users supports the notion that the transbuccal route of use can be an effective and reinforcing route of administration. The similarity in pharmacological activity and effect profile to other chewable ethnopsychopharmacological agents (tobacco, khat, pituri, and cocoa) is also consistent with areca having dependence potential.

There have however been relatively few studies assessing dependence among areca using populations. Most have been small case series, though in recent years some larger population-based studies have been conducted. Most studies indicate that the common symptoms of dependence reported by users include continued use despite knowledge of harm, craving, difficulty in controlling use, and unsuccessful attempts to cut down. Many studies have been confounded by the concurrent use of tobacco by the study population. Severity of dependence and level of use is heavier in those who also use tobacco (whether chewed and/or smoked) and alcohol. Studies assessing areca-only chewers tend to identify a lower frequency of dependence. One recent Taiwanese study reported that 44% of chewers endorsed at least one of the following items: continued use despite oral health problems, difficulty in control use or craving during periods of abstinence or difficultly in controlling use but less than 15% reported a more complete constellation of symptoms. In raising awareness of the harms associated with its use, public health officials do need to consider the potential harms of labeling a culturally sanctioned substance in terms of addiction. Such labeling may meet with considerable resistance, especially if such attention resulted in stigma, a compromised supply, or an increase in cost.

ASSESSMENT OF USE AND DEPENDENCE

Assessment of areca nut use follows the same basic principles as for any substance. Core elements include the determination of the preparation used, amount, and frequency of use and the concurrent use of other substances in particular tobacco in any form, and alcohol. The function of its use and concerns that individuals may have with regard to stopping should be explored. Dependent use will be suggested by heavier, daily use, commencement of use soon after waking, and consumption throughout the day with some users reporting queuing outside pan shops, waiting for them to open, and never being without a quid in their mouth. A more detailed assessment of use may include how many whole nuts are consumed per day, whether use is confined to certain occasions or just after meals, how long each quid or equivalent unit of dosing is chewed for and kept in the mouth and whether juices are swallowed or spat out. Since risk is dose and preparation dependent such information can form the basis of some simple harm reduction interventions (e.g. chew fewer quids per day, keep each quid in the mouth for less time, avoid combined use with tobacco products). Direct questioning as to the experience of any withdrawal symptoms on cessation of use (including an increase in tobacco smoking among those who chew and smoke tobacco) are important parts of the assessment (see later section on screening and brief interventions). Screening tools such the Self-report Screening Test for Areca Quid Abuser modified from DSM and developed in Taiwan, may be useful but require further evaluation within community settings and other countries.

The available, albeit small, literature suggests that withdrawal symptoms are seen in some users. Symptoms include mood swings, anxiety, irritability, poor concentration, sleep disturbance, lethargy, muscular aches, and an increase in appetite. They are more common and pronounced in those who also use tobacco. Many users report unsuccessful quit attempts and it is probable that the avoidance of withdrawal will be a significant driver for continued use and a common cause for unsuccessful quit attempts

CULTIVATION, COMMERCIALIZATION, AND REGULATION

In many countries, cultivation and commercialization of the areca palm is unregulated and often untaxed. Although employing large numbers of farmers, distributors, and sellers, the impact of unregulated cultivation has led to deforestation and mudslides in some areas. The wider socioeconomic impact of such unregulated and expanding production at the expense of other cash edible crops has attracted little attention. Given the profitability of nut growing, governments may need to provide compensation to farmers for growing alternative cash crops.

The commercial production, global distribution, and promotion of freeze-dried nonperishable products have widened the potential population of users. Products such as gutkha and pan masala represent a booming industry, utilizing marketing and promotional strategies frequently employed by the tobacco and alcohol companies. Advertising, colocation, and co-branding remain a problem. The provision of free packets of chewing tobacco with pan masala (where the combined product

has been banned) shows a determined marketing strategy to maintain levels of use. Areca is not controlled in the United States, widely available in the United Kingdom and Europe and is unlikely to be subject to further regulation. It is possible that should a more potent, purified version with enhanced abuse liability be developed (as occurred with methcathinone derived from khat) and commercially marketed then such legislation may need to be considered.

INTERVENTIONS

Despite its prevalence and associated health consequences there is an insufficient research to support the promotion of evidence-based guidelines regarding the prevention and treatment of areca nut abuse. Given the diverse populations and cultures in which areca is used, flexibility and local assessment of such approaches will be required in order to determine the most effective approach within each community or region.

The key strategic outcomes for areca are likely to be the same as for any drug: (1) reduction in the rate of uptake (initiation into use) and/or a delay in the age of onset of use; (2) reduction in overall prevalence of use; (3) reduction in the harms associated with its use; and (4) the early detection of problems related to use and referral for treatment. Public health and policy strategies previously adopted and evaluated for alcohol and tobacco are a reasonable starting place for policy makers considering implementing effective public health interventions. Given the co-occurrence of all three behaviors and the synergistic harm between them, linked public health and policy strategies will not optimize resources utilization but also health outcomes.

TABLE 87.5 Potential Public Health and Policy Strategies to Reduce Prevalence and Harms

Target	Intervention
Population level consumption	Interventions aimed at reducing tobacco and alcohol use. Taxation, reduction of number of outlets/opening hours
Public health education	Health awareness campaigns – use of role models Culturally sensitive, locally resourced campaigns Social marketing targeting youth population Restrictions on promotion/marketing Accurate labeling of contents and harms
Screening and early referral	Routine assessment of areca nut use and smokeless tobacco use by dental health and other primary care providers – coupled with opportunistic brief interventions and referral to specialist oral pathology or tobacco cessations service. Areca-flavored gum, nicotine gum

PUBLIC HEALTH AND POLICY APPROACHES (SEE TABLE 87.5)

Areca nut consumption is a leading preventable cause of oral malignancy. Despite its widespread use, community awareness of the risk in most using populations is low, with tobacco being the only correctly identified risk factor associated with quid chewing by subjects in one study. With research indicating that onset of tobacco and alcohol use is associated with increased rates of adolescent onset chewing, youth-focused interventions addressing all three substances appear warranted. Restriction of sales to youths and school-based projects may also act as an indirect route to educating and reducing parental use of areca (see Fig. 87.6). Other high-risk populations such as truck drivers, older women, and the military could also be targeted. Campaigns will need to take into account issues of literacy and social conformity within more remote areas.

Evidence from Taiwan supports the role of centrally regulated taxation as effective strategy to reduce the prevalence of areca use with an additional notable finding being that increasing tax on tobacco products has a significant effect on reducing rates of areca use and to a lesser extent alcohol. Given the socioeconomic spread of use within some cultures as with tobacco, it is possible that taxation will hit those with already high levels of ill health compounded by poor access to good health care and nutrition.

TREATMENT RESPONSES (SEE TABLE 87.6)

Apart from family interventions there has little work done exploring the effectiveness of more individual-based approaches such as brief intervention and substitution therapy. Given the existing evidence to support the efficacy of behavioral interventions for tobacco use conducted by oral health professionals as part of an oral health screen by dental practitioners, adapted brief interventions and opportunistic screening within primary medical and dental health care services should be developed and evaluated. A recent UK study suggests that increased training in the skills required to provide opportunistic screening and brief interventions are needed for dentists (and probably other health care workers) in high-risk areas. Addressing tobacco

FIGURE 87.6 Public health awareness campaign, Taiwan. *Copyright Adam R Winstock*

smoking independently may also have a significant effect upon areca chewing with longitudinal studies from Taiwan confirming more than twice the quit rates among chewers who had quit smoking than those who had not. Addressing concurrent tobacco use in any form should also form one of the core approaches to harm reduction. Others should be based upon reducing the frequency and duration of exposure. Encouraging use without tobacco could also be supported by the appropriate use of nicotine replacement (probably gum). To date the research conducted into smokeless tobacco prevention and treatment has yet to embrace the challenge posed by areca products.

Even where tobacco is not used a behavioral approach using a substitute masticatory such as betel nut flavored chewing gum (available in Taiwan) may be of use. Advice on gradual as opposed to sudden cessation (which is likely to be associated with a less significant and milder withdrawal syndrome) should routinely be given. Only in the most severe cases could any form of symptomatic relief be considered as indicated (in the form of low dose anxiolytics and/or sedatives for a few days). Generic advice regarding sleep hygiene and avoidance of triggers should be provided.

CONCLUSION

Areca nut remains a major cause of mortality and morbidity for many millions of users around the world, primarily though its link with oral malignancy. Dependence is more common in those who use tobacco and tobacco containing products. Targeting the use of tobacco at both population and individual levels remains a key strategy. Evidence-based approaches to smoking prevention and smoking cessation are likely to significantly lower the rates of areca nut use. The significant impact of parental use of areca on the uptake of chewing combined with the frequent association of areca consumption with the use of other harmful substances especially tobacco and alcohol mean that a broad range of interventions are required. Any attempt to address areca nut usage should be conducted in tandem with strategies targeting alcohol and tobacco. Lessons from tobacco prevention, control, and treatment suggest raising public awareness through health education campaigns and enforcing regulations on marketing and product labeling are all likely to prove effective in reducing use. Prevention of early initiation through school based interventions need to be more widely evaluated and seen in the context of broader heath promotion

TABLE 87.6 Potential Individual Harm Reduction Approaches

Target	Intervention	Aim
Cessation	Advice to stop coupled with behavioral support and counseling	Reduce duration of exposure to carcinogens
Preparation	Avoid tobacco Use ripe versus unripe nut Reduced lime Genetically modified nuts	Reduce carcinogenic load
Chewing	Chew less often, for less time Do not sleep with quid in the mouth Do not swallow juices Alternate chewing quids with gum of other less toxic masticatory	Reduce total time in contact with oral mucosa
Replacement	Areca flavored gum, nicotine gum	Reduce total consumption, reduce risk behavior
Oral hygiene reduce	Rinse mouth out after use with chelating agent	Reduce carcinogenic activation
General diet	Health educational and improved access to affordable healthy food stuffs	Reduce alcohol and increase vegetables and fruit, possibly green tea with its antioxidant effects and improve oral hygiene

campaigns. A richer scientific base is required to evidence the utility of most of these approaches since at present their use can only be recommend based on their effectiveness with other substances. Research to date supports the role of increased taxation on tobacco and areca, other approaches targeting alcohol and tobacco use and family interventions. Whether there is a role for introducing genetically modified nuts with lowers levels of carcinogenicity is at present uncertain. While reducing population levels of use may be expected to result in a decline in incidence of oral cancers, countries need to monitor the impact of any policies on the uptake of other potentially more harmful substances.

SEE ALSO

Tobacco, Khat, Kava

Glossary

Aflatoxin a naturally occurring mycotoxin that is produced by many species of the fungus Aspergillus.
OSF oral submucous fibrosis.
Pan masala freeze dried prepackaged areca nut preparation without tobacco.
quid folded leaf package containing the areca nut and flavorings and often tobacco.
The gutka freeze dried prepackaged areca nut preparation with tobacco
Trismus inability to normally open the mouth.

Further Reading

Benegal, V., Rajkumar, R.P., Muralidharan, K., 2008. Does areca nut use lead to dependence? Drug and Alcohol Dependence 97 (1–2), 114–121.
Bhat, S.J., Blank, M.D., Balster, R.L., Nichter, M., Nichter, M., 2010. Areca nut dependence among chewers in a South Indian community who do not also use tobacco. Addiction 105 (7), 1303–1310.
Boucher, B.J., Mannan, N., January 2002. Metabolic effects of the consumption of Areca catechu. Addiction Biology 7 (1), 103–110.
Chen, S.H., Lee, J.M., Liu, H.H., Wang, H.C., Ye, C.Y., 2010. The cross-effects of cigarette and betel nut consumption in Taiwan: have tax increases made a difference?. First published online: August 31, 2010. Health Policy and Planning. http://dx.doi.org/10.1093/heapol/czq041.
Croucher, R., Islam, S., 2002. Socio-economic aspects of areca nut use. Addiction Biology 7 (1), 139–146.
Gupta, P.C., Warnakulasuriya, S., 2002. Global epidemiology of areca nut usage. Addiction Biology 7, 77–83.
Sullivan, R.J., Andres, S., Otto, C., Miles, W., Kydd, R., 2007. The effects of an indigenous muscarinic drug, betel nut (Areca catechu), on the symptoms of schizophrenia: a longitudinal study in Palau, Micronesia. American Journal of Psychiatry 164, 670–673.
Trivedy, C.R., Craig, G., Warnakulasuriya, S., 2002. The oral consequences of chewing areca nut. Addiction Biology 7, 115–125.
Wang, S.C., Tsai, C.C., Huang, S.T., Hong, Y.J., 2004. Betel nut chewing: the prevalence and the intergenerational effect of parental behavior on adolescent students. Journal of Adolescent Health 34 (3), 244–249.
Warnakulasuriya, S., Trivedy, C., Peters, T.J., 2002. Areca nut use: an independent risk factor for oral cancer. BMJ 324, 799–800.
Wen, C.P., Tsai, S.P., Cheng, T.Y., Levy, D.T., Yang, H.J., Eriksen, M.P., 2005. Uncovering the relation between betel quid chewing and cigarette smoking in Taiwan. Tobacco Control 14, i16–i22.
West, R., McNeill, A., Raw, M., 2004. Smokeless tobacco cessation guidelines for health professionals in England. British Dental Journal 196 (10), 611–618.
Winstock, A.R., Trivedy, C.R., Warnakulasuriya, K.A.A.S., Peters, T.J., 2000. A dependency syndrome related to areca nut use: some medical and psychological aspects among areca nut users in the UK. Addiction Biology 5, 173–179.
Winstock, A.R., 2002. Areca nut-abuse liability, dependence and public health. Addiction Biology 7 (1), 133–138.
Yang, M.S., Lee, C.H., Chang, S.J., Chung, T.C., Tsai, E.M., Ko, A.M., Ko, Y.C., 2008. The effect of maternal betel quid exposure during pregnancy on adverse birth outcomes among aborigines in Taiwan. Drug and Alcohol Dependence 95 (1–2), 134–139.

CHAPTER

88

Khat Addiction

*Michael Odenwald**, *Axel Klein*[$], *Nasir Warfa***

*University of Konstanz, Konstanz, Germany [$]University of Kent, Canterbury, Kent, UK

**Wolfson Institute of Preventive Medicine, London, UK

OUTLINE

INTRODUCTION

Khat is referred to the young and tender leaves and shoots of the khat tree (*Catha edulis*). It is an evergreen tree of the Celastraceae family, normally reaching 6 m in height, but in an equatorial climate it might grow up to 25 m. *Catha edulis* can be found in the Abyssinian highlands, the Horn of Africa, Eastern and Southern Africa, the Arab peninsula, and Afghanistan. The khat leaves have been consumed for centuries for their mildly stimulating properties. Khat has many names including *qat* (Yemen), *jaad/chat* (Ethiopia, Somalia), *miraa* (Kenya), or *marungi* (Uganda, Rwanda). Medieval writings and legends confirm its longstanding use, but also its Janus-faced effects being described as positive and negative. Before the eighteenth century, European travelers to the region have not noticed this custom. After having been described first by the Swedish botanist Peter Forsskål (1732–1763) a high number of travelers of the nineteenth and twentieth century mentioned its growing use. Since the end of the nineteenth century with successive transport innovations, the railway network, road haulage, and air cargo, khat has made its way from a niche crop to a cash crop, today being the economic backbone for countries like Ethiopia and Yemen. More recently, mass migration of people from the Horn of Africa meant the rapid spread of khat usage to neighboring countries, Europe, and rest of the world and the development of an international market. For commercial purposes, nowadays it is grown in the highlands around the Horn of Africa, Southern Arabia, and along the Eastern African coast in altitudes of 1500–2500 m above sea level, mostly pruned to 2–7 m. Within centuries, an amazing variety of khat types have developed out of regional climatic and environmental conditions and local cultivation habits.

Soon after harvesting, the twigs and shoots are artfully rolled into bundles and wrapped into banana leaves in order to retain moisture. In most instances, these bundles are the traded units. The leaves and tender

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00088-7

stems are usually chewed and are kept in a tight wad in the cheek pocket. Within about 30 min the user experiences physiological excitability, euphoria, talkativeness, and flow of ideas known in Arabic as "mirquaan." This is followed by a quieter, more introvert phase, giving way to a gradual comedown, and often restlessness, irritability, and melancholia. Among the ethnic groups in producing regions, khat chewing has long been a regular part of social life, accompanied key social events, strengthening the social fabric. The traditional moderating khat culture restricted access to male adults and to weekends or special occasions. Traditionally, khat was chewed communally, in Yemen, Ethiopia, and Somalia in khat houses, dedicated rooms within private households often known as "mafrishes." A more regular pattern of consumption was also found among higher class users adjacent to the production region, traditionally inspiring artistic expression in architecture, handicraft, poetry, and songs. In addition, khat farmers and long-distance travelers used khat in order to stay awake and strengthen their physical abilities. It is also used during Muslim prayers and ceremonies as well as by students for being able to study all night.

PREVALENCE OF KHAT USE

The estimation of the exact prevalence of khat use is difficult, as the consumption still largely depends on socioeconomic, ethnic, and geographical factors.

In Yemen, the habit of chewing khat was once restricted to the northern highlands and to higher classes and special social happenings, while today the population at large has access and can afford to buy khat. Based on the Household Budget Survey 1998, Yemenite households spent around 9–10% of their income on buying khat. In 70% of households at least one khat chewer is reported and its use is nearly equal among all segments of the population. A recent cross-sectional study including 792 persons of the general population aged 15 or above found a lifetime prevalence of 81.6% among men and 43.3% among women; current everyday use of khat was found in 23.6% of the total sample (men 31.8%, women 8.9%). In another study, among 2500 patients of the Sana'a University dental school, 61.1% were current khat chewers – 87.0% of men and 12.9% of women.

In Ethiopia, khat chewing has traditionally been a habit in the southern and eastern part of the country, especially in Harerghe and among the Muslim populations. There are no prevalence studies for the general population on a national level. A household survey in the khat-producing, predominantly Muslim, rural community of Butajira including 10 468 adults above the age of 15 found 75% of men and 35% of women as current khat users, while daily use was found in 8.7%. In another household survey among 1028 adult respondents in the rural district of Adami Tullu, in southern central Ethiopia, with mixed-religions community, current khat chewing was reported in 40.4% of men, 18.2% of women, while daily use was found in 5.7% of the total sample. Khat chewing was also more frequent among the Oromo ethnic group, among Muslims, married respondents, those with lower education, and the farmers. A representative national assessment of 16 606 adolescents and young adults (15–24 years of age) found a 4-week prevalence of 16.2% and everyday use by 4.3%; khat chewing was more frequent among out-of-school youth. Other studies among Ethiopian high school or university students revealed current prevalence rates between 17.5 and 64.9% and differences between urban and rural areas.

In Kenya, khat chewing is traditionally practiced by some Meru tribes (Nyambene Hills) and Muslims, mostly Somali of origin, in the northern part of the country. No prevalence data for the general population are available. Some studies of patients of general health services in different parts of the country disclosed a lifetime prevalence rate of 10.7% in a region without khat production and a current prevalence rate of 29% in a khat-producing region. This was confirmed by a recent study.

Thirty years ago, a cross-sectional assessment of khat use in Somalia reported that in the north of the country, 64% of adult males from the general population regularly consume khat compared to 21% in the south. In Hargeisa, north-western Somalia, we recently found that 31.3% of males above the age of 12 chewed khat in the week before the interview. However, in a recent study among 8124 active combatants, self-reported khat use in the past week was 26.2% in the north and 50.7% in the south of the country, and excessive khat use patterns were seven times higher in the south. Press articles state that currently in Somalia per day 300 000 US$ are spent on khat and a debate is going on whether khat imports from neighboring countries drain Somalia's economy.

Other countries where khat is traditionally used include Djibouti, Saudi Arabia, Tanzania, Madagascar, and South Africa. The general trend in traditional use countries is that khat production and use spreads to other regions and segments of the population who did not have traditional contact to the leaves, for example, in Kenya, Ethiopia, Somalia, and the eastern parts of Yemen. In recent years, khat production and use has also been spreading to African countries where it has been unknown before, for example, Uganda and Rwanda. In summary, khat has become a large regional market without the involvement of multinationals. Today, it is a source of income for millions of farmers, pickers, packers, traders, and an everyday and leisure drug for up to 10 millions of people who are thought to use it everyday.

Because of mass migration, khat use spread to high-income countries of the European community, North America, Australia, and Israel. Khat use in Western countries is still limited to immigrant groups. Some studies ever investigated the prevalence of khat use among adult immigrants in Western countries, mainly among Somalis in the United Kingdom. In a nonrepresentative sample of 207 Somalis (male 152, female 55) living in London, Paul Griffiths found 78% (79% of males and 76% of females) with a lifetime history of khat use, 67% had been using it in the week before the interview, and 6% on a daily base. Another study by Shilpa Patel and colleagues found among 602 Somalis (324 male, 278 female) in four cities in the United Kingdom that 38% (231) had ever chewed khat (male 58%, female 16%), 34% had been using it in the month prior to the interview, and 3% would use it currently on a daily base. Among the 180 subjects (91 male, 89 female) of a representative sample of the Somali community in Greenwich assessed by Kamaldeep Bhui and colleagues, a current prevalence of 62.6% among male and 16.9% among females was found.

Evidence exists that khat-chewing habits have also qualitatively changed in the course of the last decades: what once has been a formalized and strongly regulated social habit, now carries features of excessiveness, informality, and decoupling from normative control in some user groups. This is manifest in the consumption of higher quantities of the drug by individuals, longer continuous consumption time, and parallel use of other drugs, for example, benzodiazepines or alcohol compared to traditional use patterns. While khat chewers used to be traditionally "initiated" at around 20 years of age, nowadays, they start using the drug earlier in life and consumption has become part of the youth culture. Furthermore, the formerly adult male habit is now practiced more and more by women and is also reported in pregnant and breastfeeding women.

IS KHAT A DRUG?

The controversy around khat and especially the question whether it is a drug is probably as old as use itself. It has been condemned by the Ethiopian orthodox church and the more fundamentalist Islamic schools of thought particularly in Saudi Arabia, who consider it "haram." Islamic scholars in Yemen, Somalia, and Ethiopia by contrast integrated khat use into religious life, including the study of the Holy Koran or to enhance religious experience as practiced by Sufi mystics. During the colonial era, arguments about moral degeneration, falling economic productivity, and the association with political unrest motivated official and largely unsuccessful bans on khat. A strong pro-khat movement prevented in most countries legal restriction. The economic advantages of khat to farmers, traders, and others involved in the khat business are evident. Furthermore, the discourse on khat as drug is misused in a context of interethnic and regional tensions, religious arguments, and in the conflict between traditional and Western parts of society.

The scientific discourse on khat shows similar fissures, usually along the lines of academic discipline: medical and pharmacological research approached khat with the underlying assumption that it is analogous to other psychoactive drugs; consequently, for more than a century research focused on the discovery of the addictive substance in the leaves. And after being first isolated by World Health Organization (WHO) laboratories during the end of the seventies, more than a decade of pharmacological research focused on the laboratory-based study of cathinone, the main psychoactive principle in the khat leaves. Only recently, a more comprehensive approach is used, for example, by studying the effects of khat leaves and by involving human subjects.

Anthropologists and social scientists, by contrast, stressed the cultural functions and traditional values. They often argue against the medical position, because the physical effects cannot explain the functions of its use and because addiction is a Western concept that does not fully grasp the case of khat. In Yemen, the social life of men develops around institutionalized khat-chewing sessions. This context to participate in business, family, and political decisions as well as in social life makes it difficult to abstain from khat. Also in the Somali tradition, the social act of khat chewing has an important function in the reinforcement of group cohesion and as a source of social support. Studies among Somali refugees living in Western countries report khat use in order to maintain identity and cope with the psychological problems associated with staying in a foreign and hostile environment or with posttraumatic experiences.

For long, researchers from different disciplines remained in their isolated silos and found no opportunity to exchange information and develop a common position. Hence, it was only in 1984 that the first international, interdisciplinary conference on khat could be organized. A wealth of new research has been produced across different fields, and there is a growing awareness of the need for interdisciplinary perspectives on so complex a topic. But the WHO expert committee and other organizations who have reviewed khat recently note that high-quality studies are still rare, for example, with human subjects and khat leaves instead of pure cathinone.

In summary, the question whether khat is a *drug* versus a part of *culture* or *tradition* should be framed differently. Today, it is clear that what is colloquially labeled a "psychoactive effect" is a complex interaction

of pharmacological effects, environmental conditions, and the users own learning experience (*see* Neurobiological Mechanisms of Drug Addiction: An Introduction). A comprehensive model of khat addiction is still a challenge for interdisciplinary research.

PHARMACOLOGY

WHO laboratories succeeded in the 1970s in isolating the alkaloid S-cathinone ([-]-2-aminopropiophenone or S-[-]-2-amino-1-phenyl-1-propanone) as the main psychoactive agent in the leaves. Other psychostimulant principles are cathine (1S, 2S-norpseudoephedrine) and norephedrine. However, the full range of compounds, including the cathedulins and their pharmacological mechanisms remain largely unstudied. Cathinone particularly resembles amphetamine in chemical structure as well as in biochemical and behavioral effects, though with about half the potency. Its pharmacological effects are considered to be amphetamine- and cocaine-like.

As cathinone is highly unstable and decomposes within 72 h of harvesting, khat leaves are preferably consumed when fresh. This historically restricted the area of consumption and markets. Cathinone content revealed to be different according to freshness of plant material and origin and correlated with the market price. Although its main psychoactive constituents, cathinone and cathine as well as some of their synthetically produced derivates (e.g. methcathinone) are controlled substances under the International Convention of Psychotropic Drugs (1971), the khat leaves are not.

Pharmacokinetics

By chewing the khat leaves, cathinone is effectively extracted into the saliva and directly resorbed through the oral mucosa and in the stomach. Maximal plasma concentrations are reached for cathinone 2–2.5 h after start of session, for cathine 2.6 h; cathinone has a mean terminal elimination half-life of 1.5–4.5 h, and cathine of approximately 5 h. In different experimental khat sessions, after the ingestion of 0.6 and 0.8 g of khat leaves per kg body weight, the maximum concentrations of cathinone in plasma were 58.9 ± 18.8 ng ml^{-1} and 127 ± 53 ng ml^{-1} and for cathine 71.2 ± 13.9 ng ml^{-1} and 89 ± 49 ng ml^{-1}. Cathinone has a mean residence time of 4.5 and cathine 10.2. After 24 h, it might not be possible to detect cathinone in blood plasma, while in urine it can be detected for about a day after chewing, in hair of habitual chewers probably after a much longer period. After ingestion, cathinone is reduced to its main metabolites (-)-norephedrine and (-)-norpseudoephedrine. Cathine has been confirmed in human breast milk and in infant urine.

Pharmacodynamics and Addictive Properties

In the central nervous system, cathinone provokes a release of catecholamines, especially dopamine, from presynaptic stores, being about half as potent as amphetamine. Moreover, it inhibits its reuptake and can produce a depletion of central dopamine. It was shown to be similarly rewarding as amphetamine or cocaine in discrimination and preference studies. Similarly, like other dopamine agonists, khat use has the potential to exacerbate preexisting psychotic disorders. Excessive khat use can elicit brief psychotic disorders and schizophrenia in vulnerable individuals. However, to date khat use and dopamine-related disorders are still understudied – evidence remains weak. Cathine had a much smaller effect in the release of central catecholamines, for example, cathinone was eight times as potent in sustaining an amphetamine-induced dopamine release. Like amphetamine, cathinone also has to a minor extent a serotonin-releasing effect in the nucleus caudatus. Besides that, cathinone also affects the reuptake and inhibits the metabolism of catecholamines and serotonin. Peripherically, cathinone and cathine are equipotent concerning the release of noradrenaline at the presynaptic storage sites and, thus, both have sympathomimetic effects.

In humans, khat effects were studied using either repeated physiological and questionnaire-based measurement of immediate reactions to khat administration or retrospective inquiry of usual reactions. Immediate euphoric effects followed by depression were demonstrated in controlled application studies. Thus, it was argued that the depressive phase after the end of the khat session motivates to continue chewing.

Peter Kalix in authoritative review of the first 15 years of cathinone research postulated a higher potential to induce dependence compared to amphetamines because: (1) cathinone has a more rapid onset of action in discrimination experiments; (2) tolerance to cathinone's anorectic effects develops faster and is more pronounced; (3) in conditioned taste aversion experiments cathinone is less aversive than amphetamine; and (4) in self-administration experiments respond rates are higher than those maintained by amphetamine.

KHAT ABUSE AND DEPENDENCY

There is an ongoing debate as to whether khat produces a physical dependency syndrome. For many years, the dominant view in the literature was that khat produces no physical dependency. This view is repeated in the recent reviews despite of the changing use patterns and the missing empirical data. However, in contrast to the study of its alkaloids, the khat leaves

have been rarely subject to studies on addiction-related processes. According to the World Drug Reports, khat is considered the main drug of abuse besides alcohol in Ethiopia and for other countries of the region. Like the case with any other drug, not every khat user will develop a dependence syndrome as defined by ICD 10 or DSM IV. Assuming a parallel between khat and amphetamine, the high prevalence of khat chewing in countries like Yemen or Somalia would produce khat dependence as high as 5–10% of the adult male population. An Ethiopian study, using the Composite International Diagnostic Interview, found a prevalence of khat dependence according to the criteria of ICD-10 of 5% among males (among females 1.3%) in representative sample drawn in a traditional khat-producing area. Khat dependence was more frequent among Muslims and associated with being economically disadvantaged.

Observational data confirm the existence of a specific "drug language" among Somali khat users: "xaraaro" means feelings of craving and nervousness, which are experienced by habitual chewers at the time of day before their usual khat intake starts; "dubaab" refers to unpleasant dreams involving the sensation of being suffocated that are usually experienced by heavy chewers in the first days after cessation. The phenomenon of "jibane" involves the use of khat in a group setting in order to reduce aversive symptoms in the morning.

Khat Withdrawal

The expert opinion on withdrawal symptoms upon discontinuation is that they are expected to be mild and only after prolonged use. However, looking at the existing empirical evidence, only very few studies present any data. Khat-related withdrawal symptoms were described in a study done between 1974 and 1976 by Kennedy and colleagues. The symptoms include profound lassitude; anergia; difficulty in initiating normal activity; slight trembling, several days after ceasing; nightmares, often paranoid in nature; for example, being attacked, strangled, or followed; the authors stated that these symptoms occur only after heavy chewing, starting a day after cessation and last several days. But they also state that heavy khat users often try in vain to stop using it or to reduce their amount of consumption. As khat chewing has become more prevalent and use patterns have profoundly changed throughout the last decades, this description might not necessarily be valid any more and needs further studying.

Khat Tolerance

Today, it is generally believed that khat use does not induce tolerance, going back to an expert evaluation by Eddy (1965). Among stimulant users, tolerance development, the upward shift in the set point for reward and the subsequent dysphoria ("opponent process") are closely related to the development of "binge" consumption patterns: users need to increase the dose and the frequency of drug administration in order to experience the desired psychological effects. For khat, it has been argued that the chewing mode of ingestion limits the possible amount to consume in a certain time and, thus, tolerance development is prevented. However, no studies have ever directly targeted the topic of tolerance to desired psychological effects, for example euphoria. Due to the physical limits of chewing the leaves, tolerance development might not only include increases in the amount of consumption per time unit but rather to the extension of the time spent for consuming khat that leads to an increase of the absolute amount ingested. Personal observations in Somalia and data from the UK Somali community indicate that heavy khat users can consume without a break for 48 h and even more without sleep. Several studies reported that such "binge" consumption patterns among khat users do occur and that experienced users are trained to chew such large quantities a novice would never manage. Other studies reported details on excessive and prolonged intake above what is considered normal, for example, that 12–14% of subjects reported a consumption of four or more bundles on one occasion and one single subject reported to consume 15 "bundles" in one day. In a study from Germany, blood and urine levels of cathine in subjects arrested for khat-related traffic offenses ultrapassed levels ever measured in controlled studies. In addition, the development of tolerance to physiological effects of khat was reported several times. In summary, the question of khat-related tolerance needs further studying. Furthermore, research needs to study the hypothesis whether effects of khat might be potentiated through sleep and food deprivation.

Psychological Dependency

While the data on physical aspects of khat dependency are scarce, researchers early on recognized the potential of khat to induce psychological dependency. This is best illustrated in descriptions of typical scenes of inner city khat markets at the Horn of Africa or in Eastern Africa at the time just before a khat delivery arrives. At these hours of day, users speed to khat markets frequently causing traffic accidents; an aggressive and nervous atmosphere prevails until the khat trucks arrive. In contrast, khat users who visit foreign countries are said to abstain without any difficulties without replacement.

On a more scientific ground, the potential to induce psychological dependence is confirmed by a number

of British studies, using the Severity of Dependence Scale, a five-item instrument thought to measure the psychological component of dependence. This instrument has recently been adapted and validated for the study of khat addiction. About 10% of a sample of Somali khat users scored at a level comparable with a clinical population with severe heroin dependence in need for treatment. In a more recent study with Yemenite immigrants living in the United Kingdom, this figure was 39%.

Neuropsychology and Neurocognitive Deficits

A common characteristic of chronic central stimulant abuse is marked neurocognitive deficits. A recent review found that there are virtually no studies and the authors developed suggestions how to study these effects among khat users. To date, only a few empirical studies have reported on neurocognitive effects of khat use. Kattab and colleagues published a report on aircrew members of an Arabic Airline who were daily khat chewers (25), occasional chewers (39), and nonchewers (24) presented for the Standard Aviation Medical Examination and participated in a standardized neuropsychological test battery. Daily khat chewers performed worst in subtests for perceptual speed, long-term memory, visual memory, and visual perception and had a faster EEG background activity compared to occasional khat chewers and nonchewers. Duration and amount of khat use were negatively correlated with performance. Recently, several studies on neurocognitive functions revealed impairments among khat users in the Netherlands, in Kenia, and in the Yemen compared to matched nonusers in respect to response inhibition, working memory, cognitive and behavioral flexibility, as well as executive functions. Indirect support of these deficits comes from observational data on khat effects on driving in Ethiopia. Although first evidence supports the hypothesis that prolonged khat abuse can produce a typical neurocognitive deficit syndrome that is similar to that produced by amphetamines, more studies are required to confirm it.

In summary, the topic of khat addiction urgently needs further empirical studies. Current evidence supports the hypothesis that excessive and prolonged khat use can produce a dependency and neurocognitive deficit syndrome qualitatively similar to that produced by amphetamine.

KHAT AND MEDICAL PROBLEMS

Numerous somatic and mental health problems have been associated with khat use. Moderate khat use is not noxious in most users and adverse effects are commonly linked with excessive use. There is growing evidence that khat can exacerbate preexisting mental health problems, as well as trigger psychosis and aggressive behavior, particularly in predisposed individuals. While there is functional khat use among people suffering from anxiety or depression, it has also been associated to suicidal thoughts. There are many reports of severe physical harms among chronic users, but other explanations have not been systematically ruled out, for example, tobacco smoking which is frequently combined with khat use, and pesticide content in the leaves, as well as exposure to polycyclic aromatic hydrocarbons and nitrosamines. Observed negative somatic consequences potentially associated to khat use include mucosal problems (oral keratotic white lesions, oral leukoplakia, and oral squamous cell carcinoma), hypertension, cardiovascular complications, duodenal ulcers, sexual dysfunction, hepatotoxicity, and reduced birth weight of infants born to khat-chewing mothers. By the same token, the argument for possible medicinal uses has only been touched on, for example, anticarcinogenic effects.

KHAT AND SOCIAL PROBLEMS

The exponential increase in khat consumption gave rise to a controversy on social problems to consumers. But this controversy is hampered by the lack of scientific studies. It is argued that modern khat consumers do not have the protective factors developed in traditional use cultures and have little awareness of negative effects. Khat has become the everyday drug of disadvantaged people in many countries. One frequently repeated argument is that khat chewing perpetuates poverty, accelerates social fragmentation, and leads to domestic violence. Khat use has been associated with moral decay, such as sexual disinhibition and sex work. The few cross-sectional studies could show some phenomena, for example, that khat use is associated with less spendings on food in poor households in Djibouti and with unprotected sexual behavior in Ethiopia. Further studies are urgently needed.

KHAT ADDICTION TREATMENT

Until today, no scientific reports on the treatment of khat addiction have been published. However, khat addicts represent a high proportion of the patients in dependence treatment in the countries surrounding the Horn of Africa although absolute numbers are very small. In Europe and North America, first reports on khat addiction emerged, exclusively among patients with a migration background from the Horn of Africa.

Thus, in high-income countries with high numbers of immigrants from the Horn of Africa, khat-addicted patients have begun to appear in mainstream addiction treatment centers, although numbers are small; treatment participation is mostly not voluntary. A higher numbers of patients exist who have switched from khat to other substances, available in the respective country. Because access to general services is difficult, several small projects in high-income countries have tried to improve access of khat-addicted individuals to drug treatment services. In the traditional khat countries evidence-based treatments for khat addiction are currently not existent. But, there is a large market for nonevaluated interventions to assist khat-addicted individuals, ranging from Western style counseling to traditional healing.

KHAT AND REGULATIONS

First discussions over international regulations were held at the League of Nations. Since then, successive reviews by expert committees of the WHO, the UK Advisory Council on the misuse of drugs, and experts in the Netherlands and Australia have concluded that khat did not merit scheduling on medical grounds and have adopted a laissez-faire regime approach, allowing khat to be imported as a vegetable product. A number of other countries have taken a drug control approach and prohibited khat (Saudi Arabia, Canada, Tanzania, Germany, France, and most Scandinavian countries). In many countries, like Spain but also the United States, the status of khat is unclear and the enforcement of controls inconsistent. The lack of consensus has added to the complexities, with London and Amsterdam emerging as entrepots for an onward trafficking trade to other EC countries and the United States. In 2012, the Netherlands have also announced to adopt a drug control approach on khat. Over the last few years the sharp rise in khat seizures in different European countries (notably Germany, Norway, Spain, Sweden, Demark, and Switzerland) has fueled concerns about a new drug with potential for abuse in Europe and North America. For the moment, consumption is mostly limited to populations originating in the Horn of Africa, although khat now appears in the growing internet-based trade of legal highs.

CONCLUSIONS

Various somatic and psychological problems have been associated with khat use, but these remain difficult to substantiate as other explanations have not been systematically ruled out. Most adverse health effects are commonly linked with excessive use. Although the addictive properties of khat are believed to be mild, a growing number of problematic users display compulsive use patterns similar to those known from stimulant users. There is a strong need for high-quality studies to further understand khat addiction.

The dramatic widespread of khat from a culture-bound to global substance of use and misuse has raised a number of challenges for researchers and the international community. As production is concentrated in Eastern Africa and Yemen, and consumption largely restricted to African and Arab countries as well as diaspora communities, there has been little initiative at national and international level to engage with the issue. There is a strong need for the development of coherent khat policies and of regulation, monitoring, and prevention measures. In the absence of alternative policy proposals there is a strong risk that khat controls nationally and internationally apply a war on drugs framework, with potentially severe consequences for farmers, traders, and the criminalization of consumers. Given the gap between the growing number of khat users on a worldwide scale and the poor understanding of pharmacological mechanisms, evolving patterns of use, and the khat economy, a substantial research program is required. Findings should inform a regulatory framework for the distribution and consumption aimed at reducing risk for individuals, communities, and society.

SEE ALSO

Inhalants, Epidemiology of Adolescent and Young Adult Alcohol, Tobacco, and Drug Use and Misuse in the United States, Cultural Influences on Youth Alcohol and Drug Use

List of Abbreviations

WHO World Health Organization

Glossary

Haram Arab word for dirty, forbidden by the Koran.
Jaad/Chat term used in Ethiopia to refer to khat; *see* Khat.
Khat leaves of *Catha edulis* that contain several central stimulants, for example, cathinone and cathine.
Miraa term used in Kenya for khat; *see* Khat.
Marungi term used in Uganda for khat; *see* Khat.
Mirquaan Arab word for khat-induced high.
Mafrish Arab word for special room in houses where khat is chewed.
Qat term used in Yemen for khat; *see* Khat.

Further Reading

Al-Habori, M., 2005. The potential adverse effects of habitual use of *Catha edulis* (khat). Expert Opinion on Drug Safety 4, 1145–1154.

Anderson, D., Beckerleg, S., Hailu, D., Klein, A., 2007. The Khat Controversy: Stimulating the Debate on Drugs. Berg, Oxford.

Feyissa, A.M., Kelly, J.P., 2008. A review of the neuropharmacological properties of khat. Progress in Neuro-Psychopharmacology and Biological Psychiatry 32, 1147–1166.

Gebissa, E., 2004. Leaf of Allah: Khat and Agricultural Transformation in Harerge, Ethiopia 1875–1991. James Currey, Ohio University Press, Oxford, Athens.

Griffiths, P., Lopez, D., Sedefov, R., et al., 2010. Khat use and monitoring drug use in Europe: the current situation and issues for the future. Journal of Ethnopharmacology 132, 578–583.

Hoffman, R., Al'Absi, M., 2010. Khat use and neurobehavioral functions: suggestions for future studies. Journal of Ethnopharmacology 132, 554–563.

Kalix, P., 1990. Pharmacological properties of the stimulant khat. Pharmacology and Therapeutics 48, 397–416.

Kassim, S., Islam, S., Croucher, R., 2010. Validity and reliability of a severity of dependence scale for khat (SDS-khat). Journal of Ethnopharmacology 132, 570–577.

Klein, A., 2007. Khat and the creation of tradition in the Somali diaspora. In: Fountain, J., Korf, D.J. (Eds.), Drugs in society: a European perspective. Radcliffe Publishing, Oxford.

Klein, A., Beckerleg, S., Hailu, D., 2009. Regulating khat – dilemmas and opportunities for the international drug control system. International Journal of Drug Policy 20, 509–513.

Krikorian, A.D., 1984. Kat and its use: an historical perspective. Journal of Ethnopharmacology 12, 115–178.

Milanovic, B., 2008. Qat expenditures in Yemen and Djibouti: an empirical analysis. Journal of African Economies 17, 661–668.

Odenwald, M., Hinkel, H., Schauer, E., et al., 2009. Use of khat and posttraumatic stress disorder as risk factors for psychotic symptoms: a study of Somali combatants. Social Science and Medicine 69, 1040–1048.

Odenwald, M., Warfa, N., Bhui, K., Elbert, T., 2010. The stimulant khat – another door in the wall? A call for overcoming the barriers. Journal of Ethnopharmacology 132, 615–619.

Warfa, N., Klein, A., Bhui, K., et al., 2007. Khat use and mental illness: a critical review. Social Science and Medicine 65, 309–318.

Weir, S., 1985. Qat in Yemen: Consumption and Social Change. British Museum Publications Limited, London.

Relevant Websites

http://www.homeoffice.gov.uk – The British Home Office (2005), Advisory Council for the Misuse of Drugs.

http://www.emcdda.europa.eu – The European Monitoring Center for Drugs and Drug Addiction (EMCDDA).

http://www.esf.org – European Science Foundation (2009), Khat Conference Webpage.

http://www.ipc-undp.org – UNDP, IPC-IG (2007), "Should Khat be Banned? The Developmental Impact."

http://www.who.int – WHO (2006), Expert Committee on Drug Dependence.

http://www-wds.worldbank.org – The Worldbank (2007), "Khat Demand Reduction."

CHAPTER

89

Internet: Immersive Virtual Worlds

Qiaolei Jiang[*], *Xiuqin Huang*[$]

[*]Nanyang Technological University, Singapore

[$]General Hospital of Beijing Military Region, Beijing, China

OUTLINE

THE DEFINITION OF IMMERSIVE VIRTUAL WORLDS

Virtual environments, in which three-dimensional displays and interaction devices immerse the user in a synthesized world, are sometimes also referred to as virtual reality (VR). VR constitutes a three-dimensional interface that puts the interacting subject in a condition of active exchange with a world re-created via a computer. However, there is no accepted definition for VR. One important reference, the US National Research Council report *Virtual Reality: Scientific and Technical Challenges*, does not attempt a definition. Rather, characteristics of a virtual environment are given. These include a man–machine interface between human and computer, 3-D objects, objects having a spatial presence independent of the user's position, and the user manipulating objects using a variety of motor channels. VR is such a complex and challenging field, and several distinct types of systems have been developed for displaying and interacting with virtual environments. Some scholars classify VR into three categories: immersive head-mounted displays (HMDs), immersive non-HMD systems, and partially immersive tabletop systems.

VR systems enable the user to feel as if they are present in a computer-generated environment. Technologies are used to create a realistic virtual world that is as much similar to real world as possible. For example, virtual, aural, haptic, and other displays are used to immerse the user in the virtual world and block out contradictory sensory impression from the real world. Graphics rendering systems are adopted for generating detailed and realistic images of the virtual world as it changes. The users' body movements are tracked by tracking systems in order to relate these to movements in the virtual world. Database systems are needed for building and maintaining detailed models of the virtual world.

A virtual environment is a "mental model" that represents a physical environment. A major part of VR research and development focuses on creating the illusion of non-mediation. In computer mediated communication, individuals apply social rules and the literature points to five characteristics that cue the idea that one is interacting with a social actor. These are (1) language, (2) interactivity to multiple prior inputs, (3) computers

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00089-9

fulfilling social roles, (4) use of human sounding speech, and (5) the possession of a human or humanlike face. According to Singhal and Zyda, Networked Virtual Environment is distinguished by the following five common features: a shared sense of space (illusion of being located in the same place), a shared sense of presence (avatars of participants), a shared sense of time (real-time interaction possible), a way to communicate (various interaction methods), and a way to share (dynamic environment that can be interacted with).

The Internet offers its users a digital environment, which can be populated with users' replicas (cymans) and objects. Within the digital environment they are free to navigate, meet other users and manipulate data usually represented in the form of three-dimensional objects. At the same time they can use audio and video channels to interact with other users using both verbal and nonverbal forms of communication. In this way, the developed online communication interface brings together remotely located collaborating parties in a shared electronic space for their communication. Nowadays, the Internet has emerged as an important social technology used for communication between individuals and groups, and has been regarded as an essential media channel for information exchange, academic research, and entertainment.

A BRIEF HISTORICAL OUTLINE OF VR TECHNOLOGY

Immersive VR has become a frequently used buzzword, while the technology of VR is not as recent as one might suppose. The first VR system appeared early in 1962, when Morton Heilig presented the first prototype of multisensorial simulator called *Sensorama*. This prototype was simulator of a real experience (a motorcycle ride through New York), with all the main features of a modern VR system but one exception: there was no possibility of interaction by the user. Few years later, Ivan Sutherland, one of the pioneers of computer graphics, described and realized the first virtual environment display system – the first HMD in 1968. It was a visual device which can display an image in front of the user's eyes, no matter where the user may be looking. Key innovative features of Sutherland's HMD included the implementation of stereoscopic vision, the totally computer-generated visual images, the adaption of the user's view according to the head's movements and the head-referenced visual display. Sutherland is now generally acknowledged as one of the founding fathers of VR. In his classic 1965 paper *The ultimate display,* he described his ideas of immersion into computer-generated environments via new types of multimodal input and output devices.

Later, the introduction of flight simulators was one of the most important precursors to VR. The innovations were driven by the greater dangers associated with training on flying the jet fighters. By the 1970s, computer-generated graphics replaced videos and models which had been used since the Second World War. Fight simulations were operating in real time, though the graphics was primitive. In 1979, the military started to experiment fight simulation systems using HMDs. In the mid-1980s, a limited three-dimensional virtual workspace in which the user interactively manipulated three-dimensional graphical objects spatially corresponding to hand position was developed. In 1984, NASA started the VIVED project (Virtual Visual Environment Display) and later the VIEW project (Virtual Interactive Environment Workstation). The goal of both projects was to develop a multipurpose, multimode operator interface to facilitate natural interaction with complex operational tasks and to augment operator awareness of large-scale autonomous integrated systems.

VR was introduced to the general public in 1989, at two trade shows organized by AutoDesk and Visual Programming Language (VPL) Research, two companies involved with NASA projects. The term Virtual Reality was originated at that time by J. Lanier, defining it as a computer generated, interactive, and three-dimensional environment in which a person is immersed. Through the 1990s, 3-D computer models have been used increasingly on the World-Wide Web using virtual reality modeling language (VRML), and faster computers provided the key to interactivity. The richness of Web page content has improved dramatically. It is possible to include color pictures, sound, movies, 3-D animation, and interactive programs. Scientists developed advanced visualization software programs and many research centers started working on VR applications in education, medicine, industry, military training, and entertainment. Nowadays, more human activities that benefit from VR technology are constantly of wireless devices and wearable computing, which represent the new frontier of the research concerning VR interfaces. The ultimate objective of this evolution is the full immersion of the human sensorimotor channels into a vivid and global communication experience. Following this approach, virtual reality can be defined in terms of human experience as a real or simulated environment in which a perceiver experiences telepresence.

THE TECHNOLOGICAL ASPECTS OF IMMERSIVE VIRTUAL ENVIRONMENTS

Immersive virtual environments are sometimes described as a particular collection of technological

hardware. For example, many people identify VR with a collection of devices, such as a computer-generated real-time graphical display accessed by the subject through the use of some type of head-mounted and panoramic displays, tracking mechanism, and other sensory input devices.

In its most well-known incarnation, virtual environments are presented to the user via a HMD where the visual information is presented to the eyes by using cathode ray tube (CRT) or liquid crystal display (LCD) technology, and auditory information is presented using earphones. The HMD has been the primary VR visual device for much of the 1990s, and is usually fitted with position tracking devices for measuring the position and orientation of the user. An alternative interface to the HMD is the Binocular Omni-Orientation Monitor (BOOM) where the display device isn't worn on the head but mounted onto a flexible swivel arm construction so that it can be freely moved in space. Tactile and force feedback is also sometimes provided through various devices, ranging from inflatable pressure pads in data gloves or body suits to force-feedback arms or exoskeleton systems. HMD and BOOM are similar devices in that the user is fully immersed in the virtual environment and does not see his/her actual surroundings.

A second common design of immersive virtual environments is through multiple projection screens and loudspeakers placed around the user. A popular implementation of such a projection system is known as the CAVE, a recursive acronym for CAVE Automatic Virtual Environment, and a reference to *The Simile of the Cave* from Plato's *The Republic*, in which he discusses about inferring reality from projections (shadows) thrown on the wall of a cave. The CAVE consists of three stereoscopic rear-projection screens for walls and a down-projection screen for the floor. Participants entering this room-like display are surrounded by a nearly continuous virtual scene. They can wear shutter glasses in order to see the imagery in stereo, and wearing a position tracker is required to calculate and render the appropriate viewer-centered perspective.

However, this exclusive focus on technology is somewhat disappointing. Immersion does not necessarily isolate the user (like an HMD) or require a special room (like a CAVE). The latest developments in the Internet, computer animation, virtual environments, and communication networks are being configured to create revolutionary tools and systems. Based on 3-D modeling and texturing tools and 3-D navigation techniques, there are a number of Web technologies that offer 3-D and interaction. Current immersive virtual environment systems use a wide variety of programming methodologies and scene description languages to handle dynamic 2-D/3-D graphics and animation within different kinds of multimedia services. Today's VR technology provides an impressive output quality. With 3-D visualization technology, the descriptions can be highly detailed. Additional features, such as posters, chat rooms, and many others, are easy to integrate. The increase in bandwidth and communication channels can enhance the interaction. In this way, the virtual environment is presented immersively to the participant. Based on the Internet, immersive digital environments are gaining in popularity because they are more easily integrated with daily activities.

IMMERSION IN VIRTUAL WORLDS

In immersive virtual worlds, man is not simply an external observer of pictures or one who passively experiences the reality created by the computer, but may actively modify the three-dimensional world in which he is acting, in a condition of complete sensorial immersion. Immersion can be understood as a property of the technical interface and as a psychological state. As for the former, immersion is a description of overall fidelity in relation to physical reality provided by the display and interaction system. As for the latter, immersion means that users can feel emotionally and cognitively present in the situation, and feel in a similar environment to the real world, with similar level of interaction.

Riva defines the "soul" of VR as a mental experience, which gives people the feeling of being really there, close to if not equivalent to the experience of actual presence. Virtual world can be created to give a 360° view within an environment. By moving around within virtual environments and interacting with objects and actors represented there, users are able to interact with synthetic or computer-generated environments. When interacting with a virtual environment or other users, the user is no longer a mere observer of what is happening on the screen. Instead, the user feels immersed in that world and can participate in it, in spite of the fact that these worlds are spaces and objects existing only in the memory of the computer and in the user's mind.

Vividness and interactivity are key factors for users' immersive feeling. Vividness is the ability of the technology to produce rich mediated environment, and interactivity is the degree of influence users have over the form or content of the environment, and the extent to which the user feels convinced of the mutual effect that he/she and the environment have on one another. Other immersion and pictorial realism factors include spatial presence, involvement, and realness. Spatial presence refers to what extent the participant has the sensation of being a part of the virtual environment physically.

Involvement means to what extent the participant's attentional resources are directed to the virtual environment. Realness appraisal describes to what extent the participant judges the virtual environment comparable to a corresponding real environment. Although the naturalness and intuitiveness of face-to-face communication is hard to achieve, the virtual environments provide additional and novel ways to enhance the weak areas of interaction.

Realistic virtual environments not only include believable appearance and simulation of the virtual world, but also imply the natural representation of participants. Users of early VR systems would normally see a representation of their hand in the virtual world to help them interact with objects and issue navigational commands. With the advent of more powerful computers and the experience gained from working with virtual environments, it is now widely accepted that the user should be represented in greater detail. The use of realistically modeled and animated humanlike avatars becomes the theme of multiple-user immersive virtual worlds, as it provides believable appearance and realistic movements. Using this embodiment regularly, the participant has a bounded, authentic, and coherent representation in the virtual world. In addition, it is useful to have the capability to make changes to the world while the participant is immersed. By changing decoration of the body through clothes and accessories, the representation also has an emergent identity. Therefore, it is likely to increase the sense of presence of the real participants in the environment.

Immersion also can mean the degree of psychological presence in virtual environments. The two dimensions are locus, that is, whether attention is directed toward the virtual or the physical world, and sensus, which is the level of attentional arousal, on a continuum from completely unconscious to fully conscious. If built properly, there can be perceptual illusion of actually being in the virtual world. The illusion involves real-time responses of the human sensory, cognitive, and affective processing systems to objects and non-mediation when he/she fails to perceive or acknowledge the existence of a medium in his/her communication environment and responds as he/she would if the medium were not there. The first characteristic of such a satisfying virtual environment is the disappearance of mediation, a level of experience where both the virtual reality system and the external physical environment disappear from the user's phenomenal awareness. The second characteristic is the possibility of building and sharing a common ground through the interaction process. In this way, a synthetic environment may combine remote and virtual presence. Remote presence refers to presence in a virtual environment through sensory extension, where users believe that they are in contact with a real, albeit remote environment. Virtual presence is characterized by the fact that users feel present in an environment simulated by a presence medium in the form of a mental model.

METHODS TO MEASURE IMMERSION

Research efforts have been devoted to finding out a measurement tool that can capture participant's sense of immersion in virtual worlds. How do we determine the extent to which a user feels immersed in the virtual environments? Basically, there are three categories of methods commonly used, including subjective measures, behavioral measures, and physiological measures.

Subjective measures rely on self-assessment by participants themselves and include a variety of techniques from subjective rating scales or questionnaires, to self-reported breakdowns of the sense of presence, to reports, comments, and interviews, to comparison-based predictions in which the subjects' task is to evaluate the difference between real and virtual scenes. Users answer questions such as "How real did the environment seem to you?", "Was the environment like a place you visited, or a series of images presented to you?". The Witmer and Singer Presence Questionnaire, and the Slater–Usoh–Steed (SUS) questionnaire are post-immersion questionnaires of this kind. The Witmer and Singer Presence Questionnaire tries to measure it by examining factors such as control factors, sensory factors, distraction factors, and realism factors. The control factors include the ability to control the relation of sensors to the environment, the speed at which the system reacts to changes caused by the user, the amount to which a user can anticipate what can happen next in the environment, the naturalness of control over the environment, and the ability to modify the physical environment. Sensory factors examine the amounts, coherence, and consistency of information arriving from different senses, the ability to perceive self-movement through the environment and the ability to search the environment actively. Distraction factors measure the possible distractions a person may experience in a virtual environment, such as awareness of the real environment, devices used to transmit the virtual environment to the user, and the observer's willingness to focus on the virtual environment stimuli presented. Realism factors measure those factors, such as environment realism and meaningfulness, as well as disorientation when returning to the real world. The SUS questionnaire has been developed over a number of years and a number of experiments at the University College of London. The questions are based on variations on one of the three themes: sense of being in the virtual environment, the extent to which the virtual

environment becomes the dominant reality, and the extent to which the virtual environment is remembered as a place. The major disadvantage of these questionnaires is the fact that they are post-immersion.

Behavioral measures provide an alternative approach to avoid the subjective nature of questionnaires. Behavioral measures examine actions or manners exhibited by the user that are responses to objects or events in the virtual environment. The premise behind these measures is that the more a participant feels immersed in a virtual environment, the more his responses to stimuli will match those behaviors he would exhibit in an identical real environment, that are produced automatically, without conscious thought. A list of behaviors that are typically exhibited around precipices was constructed, such as taking careful baby-steps, leaning away from the pit toward the walls, testing the edge of the ledge with a foot, reaching for a virtual object, greeting avatars, turning away or closing one's eyes when presented with an anxiety-provoking scene, and startle responses, orienting response and socially conditioned response such as ducking to avoid an object. Behavioral measures are more shielded from subject bias than subjective measures. Another strength is their lack of intrusion into the virtual experience. The weakness of behavioral measures is the inability to know for a balance in a VR experience which could be from the content of the environment or from the VR system itself. Behavioral measures can be exposed to experimenter bias.

There are also numerous physiological indicators that can be measured. Physiological methods attempt to measure presence by gauging changes in the subject's cardiovascular response, muscular tension, skin temperature, skin conductance, breathing rate, and ocular movement after specific stimuli or events in the virtual environments. They are more objective than subjective measures and many behavioral measures. They are generally not consciously affected by the participants, and with limited, uniform instructions, cannot be affected by the experimenters. The three commonly examined and least intrusive are change in heart rate, change in skin conductance, and change in skin temperature. Previous studies revealed that percentage change in heart rate and skin resistance had a high level of correlation with degree of immersiveness. However, the demonstration that these stimuli are directly connected with the level of immersion in virtual environments needs more systematic research.

APPLICATIONS AND EXTENSIONS OF VR TECHNOLOGIES

Virtual environments have been applied to various areas, such as scientific visualization, physical therapy, psychotherapy, online games, 3-D chat rooms, Internet shopping malls, advertising, manufacturing, electronic kiosks, location-based entertainment, HDTV, IMAX theaters, etc. As for VR applications, the goal is to use technology to create an abstract virtual space where people and their actions are simulated and shared. As technology decreases in price, and as virtual worlds are widely available through the Internet, more and more researchers will continue to embrace this ever changing technology to develop virtual environments for more application domains and benefit the society as a whole.

One popular application of virtual environment is multiplayer 3-D game. The participants of collaborative virtual environment can use avatars to interact with the contents of the world and to communicate with one another using different media, including audio, video, graphical gestures, and text. For example, in Multi-user Dungeons (MUDs), which are textual virtual environments shared by many participants at the same time, people experience new forms of communication and identity. The interactions in contemporary massive multiplayer online role playing games (MMORPGs) are more exciting. For instance, the Half-life multiplayer modification Counter-Strike (CS) has been one of the most played action games on the Internet for quite a long time. The game is very realistic, and players really have to work in teams and think what they are doing to achieve their goals. Since CS is a quite realistic military action game, it was interesting to see how much players tried to implement real-life tactics in the game as seen from the films and TV. Scholars argue that more complex virtual human embodiments will increase the natural interaction within the environment. The more natural perception of each other increases the sense of being together and thus enhances the overall sense of shared presence in the environment.

VR has been used as a tool in the assessment or treatment of physical and psychological disorders. For instance, VR allows surgeons to visualize and plan the phases and results of surgical operations. The opportunity offered by VR technology to create interactive three-dimensional stimulus environments, within which all behavioral responding can be recorded, offers experimental psychologists options that are not available using traditional techniques. An environment can be designed where the user experiences a "drug-addict world"-based virtual psychodrama or sociodrama or as part of a psychotherapeutical treatment. VR has also been used to treat acrophobia or fear of heights. There are different types of psychological Internet applications, such as self-help guides; psychological testing and assessment; information about specific psychological services; synchronous and asynchronous support groups, discussion groups, and group counseling; and

real-time counseling through chat, web telephony, and videoconferencing.

The Internet has become an important medium for commercial applications. Today, many vendors offer their goods with Internet-based technology. Using virtual environment as a marketing medium can show its worth. Immersive virtual environments could be more effective persuasion channels than the classical advertising media. One example is modeling virtual supermarkets, which is automatic generation of virtual worlds for electronic commerce applications on the Internet. An important feature is the system's ability to generate individualized supermarkets for each customer's specific needs and demands.

Education is another important application of virtual environments. The possibilities provided by the use virtual environments, such as 3-D immersion, multiple perspective, and multisensory cues offer a number of potential benefits, including experiential and active learning; visualization and reification; learning in contexts impossible or difficult to experience in real life; motivation enhancement; collaboration fostering; adaptability; evaluation and assessment; etc. Nowadays, VR has been widely used in education and training, such as augmented reality and educational toys, VR in special needs education, military training, and medical training.

By a closer inspection of the setting in which virtual environments are applied, in fact, be they clinical, experimental, or training, it appears that participants focus on resources both internal and external to the virtual environment, which are relevant to their situated activity. The need for a situated perspective has been strongly sustained in the literature on human–computer interaction. It seems more appropriate to place participants' action in a physical and symbolic scenario, hybrid in its reality, where the virtual experience is intertwined with the experience in the external world. This general position implies that the evaluation of the virtual system should extend beyond the limits of the virtual setting per se and the logic of an individual operating on an interface. The virtual environment should be considered as the fulcrum of a larger set of converging elements and its evaluation should address the specific, contextualized sequence of action performed by users.

THE PROFOUND EFFECTS ON PEOPLE'S DAILY LIVES

The ubiquity of the Internet has become a significant feature of the urban life in modern societies. The digital communication networks have allowed people to interact with virtually anybody, anywhere in the world. With the use of computer and Internet over the last decade, various virtual technologies are spreading rapidly, calling for researchers' responsible attention to the phenomenon and to its implications for users. The impact of those characteristics (such as highly visual, immersive, and 3-D environments) on human psychology and behavior is controversial, from the perspective of personal development and psychological selection, and individual participation and integration in the cultural context.

Optimists celebrate the liberation from the limits set by our natural environment. According to them, VR indulges our dreams of an ideal, pure dimension of life, where our existence is transposed on a digital code. The conventional physical constraints can be overridden; prejudices rooted in physical differences (gender, race, age) cannot be applied any longer; and time and space cease to hamper the willingness of people to meet and reach interesting destinations. Cyberspace would have the purity of our mental life without being a solipsistic experience, ushering a utopian digital society of electronically interfaced people, where our spirit freed from the destiny of our body. Pessimists have expressed their great concerns on the advances of virtual technologies. Body and action are considered the fundamental warrants of resistance and social emancipations. Moving onto a virtual mode of existence would do nothing but dissolve this warrant. The idea is that people would be left pray, unaware, of an entertaining illusion of reality, the engagement in the world being replaced with the engagement in a fake one.

To avoid these simplifications and the assumptions they derive from, many scholars have worked at a more elaborated depiction of immersion. Some have proposed the partitioning among simple, cybernetic and experiential presence, focus, locus and sensus, personal, social and environmental presence, or according to different tasks to be performed. Some have approached immersion with other purposes than measurement, for instance to investigate the relation with place, the type of conversational patterns, the kind of embodiment, thus liberating a more lively image of this experience. Some have unearthed what has been overlooked before, such as the alternation of emersion and immersion during a session.

It is true that rich and immersive interfaces can display much more information and in a more compelling way than those non-immersive technologies. In recent years, a growing number of researchers have begun to investigate the subjective experience persons report when interacting in virtual environments. As for some participants, after the virtual experience, they remember it as having visited a place rather than just having seen images generated by a computer. For some users, the virtual environment becomes the

dominant one. For instance, those people will tend to respond to events in the virtual environment rather than in the real world.

One influential application of immersive virtual environment is the highly focused, entertaining and problem-solving oriented games domain. Online games emphasize compelling content, fast pace of action and aspects of fun, and thus children and young people are usually the special category of end users. Many recent games domain is the networked 3-D-first-person-view fighting games. The various games contain a degree of teamwork, either forced by the game plot, or arranged voluntarily by a group of players. Psychologically, it has one's attention captured, and children have comparatively high expectations of such environments. However, in the long run, it can be quite problematic for children's psychological development after excessive exposure to such computer games and online games.

People's experience shows that simulation sickness is one problem for the immersive experience. Some people experienced some form of side effects, characterized by some symptoms, including ocular problems, such as eyestrain, blurred vision and fatigue; disorientation and balance disturbances; and nausea. Females tend to be more susceptible to virtual reality side effects than males. Moreover, there is no clear distinction between VR – the reconstruction of slices of life in the machine – and the building of artificial environments, where less reference is made to the characteristics of the observed world. Virtual and mixed reality environments can produce vivid experiences and generate powerful emotions. Thus, there can be problems related to online relationship and virtual property when people interact with other individuals and with the objects in virtual worlds. In addition, the easy availability of information and social contacts without needing to leave one's own room can cause isolation, hinder the building of interpersonal relationships, and produce distortion in the quality and features of the selected information. Kraut and his colleagues examine the Internet's impact on emotional well-being. The results showed that greater use of the Internet resulted in small but statistically significant increase in depression and loneliness and decreases in social engagement.

EXCESSIVE INVOLVEMENT: INTERNET ADDICTION

One of the main advantages of the online virtual worlds is the immersive characteristics, sealing off the physical environment and presenting the senses with an inclusive virtual environment. Immersion is such a condition in which the user is isolated from the external stimuli. The separation of the psychological realm from the physical and the simplified idea of a full immersion into the simulation are at the core of common opinions on computer technologies and VR, which have gathered around extreme poles. When the immersive virtual Internet worlds bring people benefits, research suggests that excessive involvement with immersive virtual worlds may cause symptoms traditionally associated with substance-related addictions, such as uncontrolled craving for Internet activities, tolerance, and withdrawal.

The term Internet addiction disorder (IAD) was originated in the United States. It is the psychiatrist Ivan Goldberg who first coined the term Internet addiction and proposed it as a disorder in a satirical hoax in 1995, however, at that time he did not expect so many people would take it as a serious problem. The first empirical study on IAD was conducted by Young in 1996, a pioneer and the leading proponent in this field. Since Young's first work, IAD has received attention from multiple disciplines. The literature reveals IAD has grown rapidly under various names, including pathological Internet use, Internet addicted disorder, Net addiction, high Internet dependency, cyberspace addiction, online addiction, etc. Research on Internet addiction has proliferated, addressing topics as diverse as nomenclature, definition, classification, epidemiology, assessment and diagnosis, treatment and prevention.

There have been different views since the advent of Internet addiction. Firstly, whether Internet addiction is valid as a distinct disorder or it is an expression of other primary disorder is unknown. Some researcher suggests that Internet addiction appears to be a common disorder that merits inclusion in *DSM-V*. While some argue that excessive video game use may be symptomatic of other primary disorders, like depression, and/or the result of poor time management skills, rather than a bona fide addiction. Secondly, debates are focused on its nomenclature and classification accordingly. Some investigators linked it to addictive disorders, grouping it alongside alcohol and drug use disorder, so called it Internet addiction or Internet dependency. Others linked it to impulsive control disorder (ICD) and called it pathological Internet use, whereas others linked it to obsessive–compulsive disorder (OCD) and called it compulsive Internet or computer use. In addition, some scholars argued that people were not addictive to Internet itself but to Internet activities, and the Internet can be accessed not only by computer but also by cell phone, so other terms have also been offered. For instance, Griffiths believes that technological addiction, including Internet addiction, is a branch of behavioral addictions. Some researchers consider pathological electronic medium use (PEMU) is more encompassing and incorporate new electronic technologies compared with Internet

addiction. The debates are ongoing, however, a change proposed by *DSM-V* Work Group is to alter the name of substance use disorder (SUD) chapter in *DSM-IV* to "Addiction and Related Disorders," in which gambling disorder will be included as a non-substance use or behavioral addiction, and Internet addiction will be recommended for the appendix with references to the existing literature in order to encourage further research.

Internet addiction is considered to be a broad term covering a wide variety of Internet activities, usually including online gaming, online relationship (communication), cybersex, information gathering or downloading, etc. Internet addiction is characterized with excessive or uncontrolled preoccupations, urges, or behaviors regarding computer use or Internet access which leads to impairment and distress. While there are different definitions available for the phenomenon, some researchers have adapted SUD, but others use pathological gambling as reference, which result in an inconsistent definition of Internet addiction.

A number of studies have explored the etiology and risk factors for Internet addiction, including internal factors, namely personality and motivations of addicts, as well as external factors, such as Internet access and Internet activities. The personality traits of Internet addicts appear to involve introversion, neuroticism, and impulsivity, such as loneliness, shyness, avoidant, aggressive, hostile, and low-esteem. The motivations related to dysfunctional coping, socialization, and personal satisfaction seem to serve as risk factors for Internet addiction. Generally speaking, the general characteristics of Internet and/or virtual environments, such as virtual, anonymous, immersive, flow, 3-D high visual, always available, convenient, interactive, and inexpensive, make users addicted to it easily.

Contemporary multiplayer 3-D games can be considered as important applications and extensions of VR technologies for popular entertainment. It is expected that immersion and Internet gaming addiction are positively correlated because immersion leads to extended playing sessions due to sense of being lost in the game where players lose awareness of themselves and their surroundings. Studies showed that both immersion and flow (i.e. the optimum experience a person achieves when performing an activity) did correlate with Internet addiction. One study explores the relationship between time spent on various Internet applications (including downloading, social networking, chatting, online games, and casual games etc.) and compulsive Internet use (CIU) in a large sample of adolescents. The results reveal that online gaming has the strongest association with CIU. Other studies also show that individuals with Internet addiction were more likely to play online games. The latest game software sale rankings indicate that the public prefer MMORPGs and real-life simulation games. Another study of Internet addiction and MMORPG users' experience reveals that curiosity, role-playing, belonging, obligation, and reward can be used to predict addiction to MMORPGs. Moreover, researches indicate that structural characteristics of Internet games appear to be related to addiction. Specifically, online games are found to be more addictive than offline video games. In addition, particular game features such as adult content and rare in-game items are significantly more welcome among addicted players. Moreover, players with a higher tendency toward Internet addiction are usually more proud of their virtual characters (avatars). Some regard their avatars so superior that they wish to become like that in real life.

Internet addiction is associated with large amounts of time spent on Internet activities, lack of sleep, and a shortage of social interactions. Moreover, similar to any other substance-related addiction, the following six symptoms are experienced, namely salience, mood modification, conflict, withdrawal, cravings, and relapse. Additionally, addicts perceive Internet activities as providing compensation for needs which are not met in their real lives, and the Internet has become the focus of their lives.

Diagnostic instruments have been developed to screen Internet addiction. Young modified the diagnostic criteria of pathological gambling in *DSM-IV* to construct an eight-item questionnaire to diagnose pathological Internet use. According to the concepts of behavioral addiction, Griffiths proposed six IAD symptoms: salience, mood modification, tolerance, withdrawal, conflict, and relapse. Shapira et al. have also proposed diagnostic criteria for problematic Internet use according to the concepts of impulse control disorder in *DSM-IV-TR*. Shapira's three item criteria is (1) uncontrollable; (2) markedly distressing, time-consuming or resulting in social, occupational or financial difficulties; and (3) not solely present during hypomanic or manic symptoms. However, there was little empirical evidence to support their contents and cutoff point. Chinese researchers Tao and Ko have also proposed diagnostic criteria for Internet addiction respectively, and the validity and reliability have proved to be good. Subjects of the two studies were adolescents or young adults, and the subtypes of Internet addiction were not designated clearly. Similarly, some assessment instruments have been created, but few of them have undergone rigorous reliability and validity tests, and none of them wholly captures the various dimensions of Internet addiction. Still, there has not been universally acceptable gold standard assessment or diagnostic tool until now.

With these definitions and assessment methods, estimate of prevalence of Internet addiction has quite a large range from 0.9 to 38%, usually being high in Asia. The

common age range of onset of Internet addiction is from 13 to 30 years old. Some studies show that younger online-game players are generally more prone to IAD. As for gender, males show higher dependence on video games than females, which has been explained by a functional magnetic resonance imaging (fMRI) study contrasting a space-infringement game with a control task. In the study, males show greater activation and functional connectivity compared to females in the mesocorticolimbic system.

Pathophysiology studies appear to show that neither the causes nor the consequences of Internet addiction are restricted to psychosocial factors. Perhaps non-substance-related behaviors such as gambling or Internet gaming could also produce intense reward activation and lead to compulsive behavior similar to addiction. Results of fMRI studies reveal that significantly stronger activation in Internet gaming addicts compared with healthy controls has been found in the left occipital lobe, left parahippocampal gyrus, left dorsolateral prefrontal cortex, right orbitofrontal cortex, bilateral anterior cingulate, and medial frontal cortex.

Internet addiction can lead to a wide variety of negative consequences that may require professional treatment. These include psychological problems, such as inattention, decreased self-appraisement and lower psychosocial well-being; loneliness, depression, anxiety, aggression/opposition, and hostility; and decreased academic and occupational performances. Some studies demonstrated that Internet gaming addiction develops as playing times increase significantly. As loss of control, a narrow behavioral focus and serious conflicts appear. Many individuals with Internet addiction will meet the criteria for comorbid psychiatric disorders, such as anxiety, depression, attention deficiency and hyperactivity disorder (ADHD), SUD or personality disorder.

As for treatment and prevention, special organizations have emerged in many countries. Many psychosocial and pharmacological approaches have been utilized to treat Internet addiction. The psychotherapy methods include cognitive behavioral therapy, motivational interviewing, twelve-step program, etc. The formats of psychotherapy involve group, individual and family therapies, as well as the intervention based on schooling. Pharmacotherapy for Internet addiction is still quite rare. Medicine, such as bupropion, and neltrexone, used to treat SUD have been reported to treat Internet addiction. For many cases, medicine, such as antidepressants and atypical antipsychotics are applied to treat comorbidities of Internet addiction. Many treatment approaches recommended are based on clinical experience, only few on empirical data. Several approaches have been established for preventing Internet addiction from both practical and academic perspectives. One approach is using software to make the users stop gaming after excessive play. For example, the Chinese government has launched a nationwide campaign with some anti-addiction software. Once the software is installed, online game players will be forced to logout after certain hours of gaming, otherwise they may lose in-game magic, virtual items, and gaming credits. Another approach is to identify potential addicts and give warnings or appropriate education about this issue.

SUMMARY

Nowadays, computers and the Internet are an integral part of people's daily lives. Artificial three-dimensional worlds are considered to be one of the most user-friendly metaphors for human–computer interaction. People have become accustomed to comfortable, easy-to-use interactive multimedia systems, with removal of time and distance constraints. The networked virtual environments are not just animating the vastness of space beyond the screen, they are also punching out at us in true 3-D movie fashion and starting to colonize the "inner space" of our private mental models. VR systems enable the user to feel as if they are present in a computer-generated environment. Thanks to the Internet, the new society is effectively a multiplayer mixed reality game in which everyone can win, and everyone can be virtually famous, infamous, or whatever else they want to be.

As a double-edged sword, Internet has profound effects on human psychology and behavior. Initially immersive audiovisual collaborative virtual environments appeared to hold great promise as a way of supporting social interaction for both business and leisure. However, the acquisition of memes through the filter of technology cannot replace the basic role of social interactions and direct experience. One concern is that Internet use may lead to declines in social involvement, since many of the online relationships represent relatively weak ties with strangers, acquaintances, or nonintimate kin. These types of social contact online are found to provide less social support than more intimate ties do. Anonymity and possibility of concealing the participants' real identity can have bewildering effects. Therefore, Internet use has been regarded to be associated with declines in the psychological well-being, such as depression, loneliness, etc.

The violent and sexually explicit content on the Internet is another major concern among parents, educators, and policy makers. Various applications and extensions of immersive virtual worlds on the Internet, with MMORPGs as representative, involve aggression and competition, and violence is an integral part of many online games. Teenagers and young adults are particularly susceptible to the lure of the Internet, and they are described as being obsessed with the Internet. The immersive virtual worlds online present

opportunities for novel interactions, but also bring many problems, such as the risk of Internet addiction. More and more scholars from different countries have conducted studies on symptoms, diagnostic criteria, etiology, subtypes, epidemiology, pathological basis, comorbidity, treatment, prevention, and other issues related to Internet addiction.

SEE ALSO

Internet Addiction: Cybersex, Video Game Addiction, Shopping Addiction, Behavioral Economic Factors in Addictive Processes, The Cell Phone in the Twenty-first century: A Risk for Addiction or a Necessary Tool?, Overuse of Social Networking

List of Abbreviations

BOOM	Binocular Omni-Orientation Monitor
CIU	compulsive Internet use
CS	counter-strike
fMRI	functional magnetic resonance imaging
HMDs	head-mounted displays
IAD	Internet addiction disorder
MMORPGs	massive multiplayer online role playing games
SUD	substance use disorder
SUS	Slater–Usoh-Steed

Further Reading

Griffiths, M.D., 2010. Online gaming addiction: fact or fiction? In: Kaminski, W., Lorber, M. (Eds.), Clash of Realities, pp. 191–203. Munch: Kopaed.

Haraway, D.J., 1991. Simians, Cyborgs and Women: The Re-invention of Nature. Free Association Books, London.

Hsu, S.H., Wen, M.H., Wu, M.C., 2009. Exploring user experiences as predictors of MMORPG addiction. Computers and Education 53 (3), 990–999.

Huang, X., Li, M., Tao, R., 2010. Treatment of Internet addiction. Current Psychiatry Reports 12 (5), 462–470.

Ko, C.H., Liu, G.C., Hsiao, S.M., Yen, J.Y., Yang, M.J., Lin, W.C., et al., 2009. Brain activities associated with gaming urge of online gaming addiction. Journal of Psychiatric Research 43 (7), 739–747.

O'Brien, C.P., 2010. Commentary on Tao et al. (2010): Internet addiction and DSM-V. Addiction 105, 565.

Riva, G., Davide, F., IJsselsteijn, W.A., 2003. Being There: Concepts, Effects and Measurements of User Presence in Synthetic Environments. IOS Press, The Netherlands.

Riva, G., Galimberti, C., 2001. Towards Cyberpsychology: Mind, Cognition and Society in the Internet Age. IOS Press, Amsterdam, The Netherlands.

Seah, M., Cairns, P., 2008. From immersion to addiction in videogames. In: England, D., Beale, R. (Eds.), Proc. of HCI, vol. 1. BCS, pp. 55–63.

Slater, M., Wilbur, S., 1997. A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. Presence: Teleoperators and Virtual Environments 6, 603–616.

Smahel, D., Blinka, L., Ledabyl, O., 2008. Playing MMORPGs: connections between addiction and identifying with a character. Cyberpsychology and Behavior 11 (6), 715–718.

Snodgrass, J.G., Lacy, M.G., Francois Dengah, H.J., Fagan, J., Most, D.E., 2011. Magical flight and monstrous stress: technologies of absorption and mental wellness Azeroth. Culture, Medicine and Psychiatry 35 (1), 26–62.

Stanney, K.M., 2002. Handbook of Virtual Environments: Design, Implementation, and Applications. Lawrence Elbaum Associates, Mahwah, New Jersey.

Tao, R., Huang, X., Wang, J., Zhang, H., Zhang, Y., Li, M., 2010. Proposed diagnostic criteria for Internet addiction. Addiction 105, 556–564.

Vince, J., Earnshaw, R., 1998. Virtual Worlds on the Internet. IEEE Computer Society, Los Alamitos, California.

Relevant Websites

1. http://www.netaddiction.com/ – Center for Online and Internet Addiction.
2. https://wiki.duke.edu/display/DIVWP/Home – Duke Immersive Virtual World Projects Wiki.
3. http://ovrt.nist.gov/hotvr.html – Hot virtual reality sites.
4. http://conferences.computer.org/vr/2011/ – IEEE virtual reality.
5. http://www.virtual.gmu.edu/ss_worlds/index.htm – Science Space's immersive virtual worlds.
6. http://www.bilawchuk.com/mark/ – The history of virtual reality, the technology of virtual reality, and its projected social implications.

CHAPTER

90

Waterpipe Smoking

Wasim Maziak[*,$]

[*]School of Public Health, University of Memphis, Memphis, TN, USA
[$]Syrian Center for Tobacco Studies, Memphis, TN, USA

OUTLINE

INTRODUCTION

The Waterpipe

Only about a decade ago, including questions about the waterpipe (a.k.a. hookah, shisha, narghile, arghile; Fig. 90.1) in epidemiological studies of tobacco use among youth would have seemed unwarranted, even in the Eastern Mediterranean (EM) region considered by many as the cradle of this tobacco use method. Not including such questions nowadays, no matter where the study is conducted, will likely be a serious flaw. This reflects the dramatic changes in youth's tobacco use patterns worldwide, with non-cigarette forms, led by the waterpipe, becoming increasingly popular. The waterpipe usually consists of a fired-clay head, metal body, glass water bowl, and leather or plastic hose (see Fig. 90.1). But while waterpipe's shape and composition can differ from one society to another, there is a great similarity in how waterpipes are currently used by youth around the globe.

In the most common form of waterpipe use, burned charcoal pieces are placed on top of a perforated

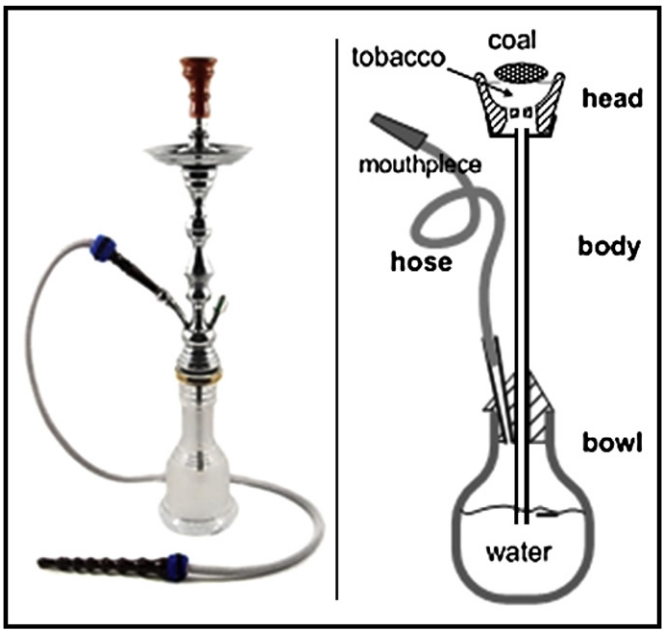


FIGURE 90.1 Actual waterpipe (left) and schematic (right) showing main parts. *Reproduced from (Maziak, 2008, 2011).*

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00090-5

aluminum foil separating it from the flavored tobacco mixture in the head. When the smoker breathes in from the hose's mouthpiece, charcoal-heated air is drawn in to become smoke as it passes through the tobacco mixture, and cools as it bubbles through the water before inhalation by the smoker (Fig. 90.1). The passage of smoke through water prior to inhalation underlies much of the common misperception of waterpipe's reduced harm (due to assumed "filtering" effect of water). The most popular type of waterpipe tobacco used today is called *Maassel* (a.k.a. shisha tobacco), which is a wet mixture of tobacco, sweeteners (e.g. molasses), and flavorings. Waterpipe charcoal products range from traditional earthen kiln charcoal to quick-lighting products that are particularly common in the United States and western societies. While evidence about waterpipe's harmful and addictive properties is emerging, treatment and policy interventions to curb waterpipe use continue to lag behind. This chapter provides an overview of the global waterpipe epidemic and the evidence about its harmful and addictive properties as it pertains to the development of treatment and policy interventions to curb waterpipe spread.

THE EPIDEMIOLOGY OF WATERPIPE SMOKING

Factors Contributing to the Global Re-emergence of Waterpipe Smoking

In recent years, waterpipe use has been witnessing a surge in popularity, especially among youth in the EM region. Beginning in the 1990s, what seemed to be a tobacco use method destined for extinction witnessed a global revival that was sparked in the EM region. The period of revival of the waterpipe coincided with the introduction of manufactured *Maassel* and the explosive increase of its available flavors. While the apple *Maassel* was somehow the predominant flavor early on, current variety include strawberry, grape, anise, apricot, cappuccino, coke, banana, passion fruit, peach, berries, papaya, mint, mango, coffee, candy, chocolate, vanilla, walnut and the list goes on. It is obvious from this list that many of those flavors are created with marketing to young people in mind.

The widespread reduced-harm perception of waterpipe is perhaps contributing to waterpipe use among youth, especially the more health conscious. For example, studies among future health professionals show that current waterpipe use is practiced by 20.6% of medical students in Lebanon, 23.5% of medical students in Syria, and 28.6% of medical students in Turkey. A recent study of primary health care providers in Aleppo-Syria shows not only the widespread of waterpipe smoking in this population (e.g. 24% of male physicians smoked waterpipe), but that health care providers' smoking represents a barrier for the implementation of smoking cessation practices in primary health care.

The aromatic and mild smoke of *Maassel*, its wide variety and availability, and its simplification of the waterpipe preparation process were perhaps critical for the renewed emergence of the waterpipe. Before the *Maassel* era, waterpipe smokers mostly used more raw forms of tobacco that needed some sort of processing and manipulation before being applied to the waterpipe (e.g. addition of water, crushing and pressing within a piece of cloth). At the same time, the Internet and other transnational media (e.g. satellite TV) commercialized and glamorized this practice, particularly among youth. In sum, manufactured *Maassel*, the reduced-harm perception, the thriving café culture, and mass media have perhaps created conditions for a perfect storm that sparked the global waterpipe epidemic.

The Spread of Waterpipe Smoking Among Youth

The global spread of waterpipe smoking among youth is backed by recent surveys from around the world. For example, results of the Global Youth Tobacco Survey (GYTS) involving data from 20 countries and the Gaza Strip-West bank in the EM suggest that current (past month) waterpipe smoking (12.9%) is more than double that of cigarette smoking (5%) among 13–15-year olds (Fig. 90.2). In Karachi Pakistan, ever waterpipe use was reported by 27% of school students ($n = 646$, mean age 15 years) and 54% of university students ($n = 450$, mean age 21 years), while current waterpipe

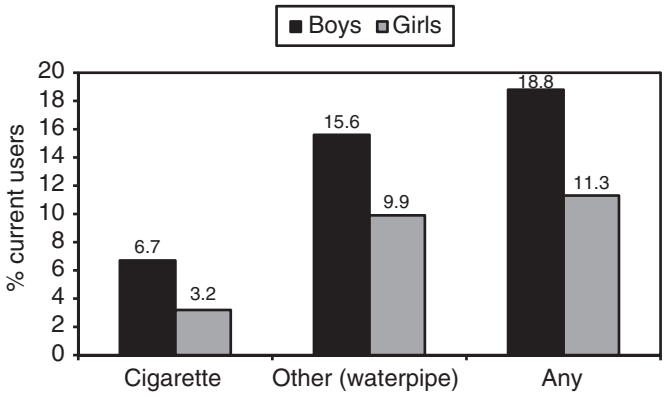


FIGURE 90.2 Data from Global Youth Tobacco Survey in the Eastern Mediterranean (92.075 students in 20 countries) showing how other tobacco use (mainly waterpipe) is becoming more widespread among 13–15-year olds in the region *(constructed based on Warren et al., Lancet, 2006).*

use was reported by 17% of school and 33% of university students. Among Jordanian university students ($n = 548$, mean age was 21.7 years), a recent study shows an amazing 42.7% prevalence of current waterpipe smoking. A recent survey of 2972 adult Kuwaitis showed that 45% smoked only waterpipe, 12% only cigarettes, and 8% both waterpipe and cigarettes.

Outside of the EM, where until recently tobacco surveys have rarely assessed waterpipe use, a new picture is emerging. Ever waterpipe use for example, has been reported by 38% of a sample of British university students ($n = 937$, mean age 20 years), and 40% of a sample of French high school students ($n = 920$, mean age 18 years) experimented with other tobacco products including the WP. Studies in US universities found that current waterpipe use ranges from 7 to 20%. For example, among a sample of 8745 students in eight colleges in the United States, 29.5% reported ever waterpipe use, and 7.2% reported current waterpipe use. As Fig. 90.3 shows, waterpipe use in US universities is approaching that of the EM, and is extending beyond any specific ethnic denominator. The spread of waterpipe smoking among US college students is also reflected in the dramatic rise (400% since 1999) of hookah lounges near US campuses.

These trends are not confined to small-scale surveys, but are beginning to show in state and national surveys. In a 2006 national survey of Estonian students ($n = 13\ 826$, age 11–15 years), waterpipe use was reported by 25% of boys and 16% of girls. Statewide youth tobacco use surveys in the United States indicate that 7% of 12th graders were current waterpipe users in Arizona, while 11% of high school students in Florida were ever users. According to the 2008 New Jersey Youth Tobacco Survey involving a representative sample of 3010 high school students, 9.7% of high school students are current waterpipe users. In Canada, the 2006 Canadian Tobacco Use Monitoring Survey shows that 8% of youth aged 15–24 years had ever used the waterpipe. The available evidence from Australia comes from a telephone survey conducted in 2004 and involved 1102 Arabic-speaking residents in south-west Sydney, where current waterpipe smoking was reported by 11.4% of participants.

Yet the most compelling evidence about waterpipe's growing impact on youth's tobacco use globally comes from a recent GYTS report looking at time trends of tobacco use (1999–2008) among more than half a million youth ages 13–15 years (involving 209 surveys in 95 countries and five areas). This global surveillance effort shows that while cigarette smoking is either stable or declining, other forms of tobacco use are showing a rising trend, most notably waterpipe smoking. As a result, questions about waterpipe are becoming an essential part of tobacco use surveys worldwide (e.g. waterpipe questions will be added to the 2010 Monitoring the Future Survey).

What began as a Middle Eastern phenomenon in the 1990s of the past century has made it quickly to the global tobacco use arena in a way that we have not witnessed perhaps since the global cigarette epidemic.

Epidemiological Patterns of Waterpipe Smoking

The salient epidemiological pattern of waterpipe smoking is the relation to age, with younger people more likely to smoke waterpipe. For example, in a population-based survey in Syria, waterpipe smoking was more than twice as popular among younger adults compared to older ones. This is quite a distinguished pattern from waterpipe smoking before the current epidemic (before 1990s), where its use was mainly seen among older men in the EM. Other noted epidemiological patterns of waterpipe use include its predominance among educated and affluent youth, within a social setting, and its relation to family waterpipe use and attitude. For example, data from within and outside the EM suggest that waterpipe smoking is favored by educated and affluent youth compared to other slices of the society (Fig. 90.4). Whether this pattern represents a transient phase of the waterpipe epidemic reminiscent of the early stages of the cigarette epidemic, or will continue beyond, remains to be seen. Undoubtedly, the continuation of strong

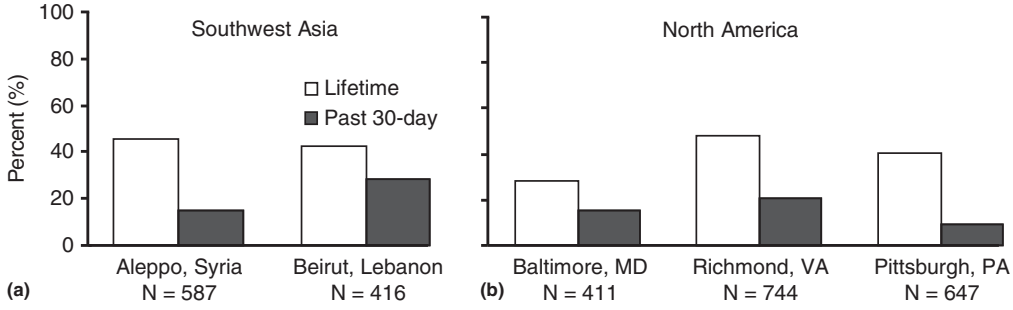


FIGURE 90.3 Prevalence of current WP smoking among college students from the Middle East and US. *Reproduced from (Cobb et al., 2010).*

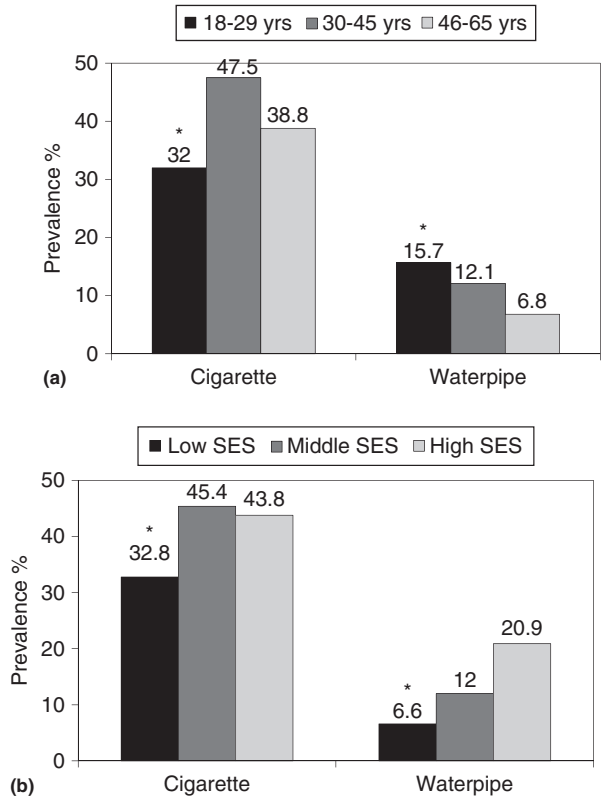


FIGURE 90.4 Data from the Aleppo Household Survey showing the clear age (a) and socioeconomic (b) relation of waterpipe smoking, with younger and more affluent people more likely to smoke waterpipe. $^{*}P < 0.05$ according to chi square test for trend. *Reproduced from (Ward et al., 2006).*

promotion and marketing of waterpipe through the Internet, the targeting of college campuses by waterpipe venues (e.g. hookah bars and lounges), the emerging data about waterpipe's health hazards, and the shape of policy initiatives to curb waterpipe use will likely influence the evolution of waterpipe epidemic in different societies.

Another distinctive pattern of waterpipe use, at least in the EM context, concerns girls/women's susceptibility to waterpipe smoking compared to cigarettes. In a region where cigarette and waterpipe smoking are by far the main tobacco use methods, the fact that 6.7% of boys in the GYTS (2006 global report) and 3.2% of girls smoked cigarettes, while 15.6% of boys and 9.9% of girls used other tobacco products (mainly waterpipe) demonstrates the dramatic spread of waterpipe use among youth in the EMR (Fig. 90.2). This spread of use among girls/women seems particularly related to social tolerance toward waterpipe smoking in this group compared to cigarettes. For example, in a study by Tamim and colleagues (2007) of 2443 school children in Lebanon (average age 15 years), about one quarter of cigarette smokers, compared to two-thirds of waterpipe smokers, said that one or both of their parents knew about their smoking. Unlike cigarette smoking, gender was not a predictor of waterpipe smoking status among the studied school children. In the study of adults in Kuwait alluded to earlier, women were more likely to smoke waterpipe compared to men, who were more likely to smoke cigarettes. It looks that in traditional societies of the EM, where women's behavior is more in the public eye and under scrutiny, women's use of the "traditional" waterpipe is seen as more conforming to the local culture than women smoking the "western" cigarettes.

In the long run, this will likely lead to fundamental changes in tobacco use patterns in societies of the EM marked by the decimation of the long-believed "immunity" of EM women to smoking. Once hooked on nicotine, it will be hard to imagine how women waterpipe smokers can abstain from cigarettes in situations, where access to waterpipe is not available. Because of its size and setup, waterpipe is not as accessible and portable as cigarettes, which can encourage switching between these two tobacco use methods. This is bad news for women's health, as well as for their children's health, as evidence from the EM shows that mothers' smoking is more important than fathers' as a determinant of children's exposure to secondhand smoke.

THE PUBLIC HEALTH IMPLICATIONS OF WATERPIPE SMOKING

Waterpipe's Harmful Potential

One of the main potential drives behind waterpipe popularity amongst youth worldwide lies in its perceived reduced harm compared to cigarettes. Several factors can underlie such perception including; waterpipe's predominantly intermittent use patterns, the smoothness of waterpipe smoke compared to cigarettes (waterpipe smoke is usually cooler and less irritating than cigarette's), and perhaps most importantly the assumed smoke "filtering" properties of water. Generally, the harmful/addictive profile of waterpipe use compared to cigarettes is largely under-researched and is likely to be influenced by the properties of waterpipe smoke, pattern, duration and frequency of use, and type of tobacco and charcoal used. Advances in our understanding of the full spectrum of waterpipe-induced morbidity and mortality are also hindered by the novelty of waterpipe epidemic relative to the long latency of important smoking-related health outcomes such lung cancer and cardiovascular disease.

However, several lines of evidence are converging to provide an alarming picture of waterpipe's harmful potential. For example, recent systematic reviews of the evidence concerning the health effects of waterpipe

smoking suggest that it is more than doubles the risk of lung cancer, respiratory illness, and low birth weight, and can influence lung function in manner similar to cigarettes raising the potential of later chronic obstructive pulmonary disease (COPD) development as a result of waterpipe use. According to these reviews, many of the reviewed waterpipe-health effects studies suffer from methodological shortcomings. On the molecular level recent research shows that waterpipe smoking is genotoxic to lymphocytes at a higher magnitude than that seen with regular cigarette smoking.

Awaiting high quality studies of the long-term health effects of waterpipe smoking, other lines of evidence allow for a forecast of future waterpipe morbidity and mortality. Foremost, although waterpipe use patterns are predominantly intermittent, waterpipe smokers inhale on average in a single session about 150 times the amount of smoke compared to a single cigarette (Table 90.1). Waterpipe smoke, moreover, seems far from being "filtered". Research shows that waterpipe smoke contains many of the same toxicants found in cigarette smoke (e.g. those associated with cardiovascular disease such as carbon monoxide "CO"; lung cancer such as polycyclic aromatic hydrocarbons (PAH); addiction via nicotine; Table 90.2), and these are delivered efficiently to the smoker. For example, analysis of smoke generated by waterpipe users, through direct sampling during the smoking sessions, shows that during a single use session averaging 1 h, users drew a mean of 119 l of smoke containing 150 mg of CO, 4 mg of nicotine, and 602 mg of "tar".

Measurement of users' exposure to waterpipe-related toxicants reveals that, relative to a single cigarette for example, a single waterpipe session exposes the smoker to 3–9 times the CO and 1.7 times the nicotine (Fig. 90.5). Most of the CO in the mainstream smoke of waterpipe is likely originating from the charcoal, adding a new dimension to the harmful potential of this tobacco use method. In fact, reports of waterpipe-related emergency room admissions due to acute CO intoxication are beginning to emerge in the medical literature. Exposure to these toxicants can be inferred as well from the acute physiological responses associated with waterpipe smoking including; increased heart and respiratory rates, blood pressure, pulse pressure, and a decrease in some lung function parameters (e.g. forced expiratory flow 25–75% (FEF 25–75%), peak expiratory flow rate (PEFR)).

TABLE 90.1 Mean Puff Topography for Waterpipe and Cigarette Smokers

Topography variable	Waterpipe $N = 61$	Cigarette $N = 56$
Puff number	169	12.7
Puff volume (ml)	511	48.6
Puff duration (s)	3.2	1.5
Inter-puff interval (s)	12.6	21.3
Total volume (ml)	79 100	523

Note: waterpipe topography data from Maziak et al. (2009); cigarette topography data from Djordjevic, Stellman, and Zand (2000). Note that WP smokers inhale in one puff almost as much smoke as cigarette smokers inhale for the whole cigarette.

TABLE 90.2 Machine-generated Smoke Content Using Realistic Puff Parameters for a Single Waterpipe Episode and a Single Cigarette

Toxicant (mg)	Waterpipe[1]	Cigarette[2]	Ratio
Nicotine	2.96	1.74	1.70
CO	145	22.3	6.50
Tar	802	17.3	46.36

From Cobb et al. (2010).

A lot of these harmful substances are emitted to the surrounding air putting non-smokers at risk. For example, two recent studies involving human and machine-smoked waterpipes showed that waterpipe smoking generates high levels of ambient air toxicants/carcinogens (e.g. PAH, metals, CO, NO, as well as particulate matter), that are comparable to smoking 2–10 cigarettes for a 1-h waterpipe session.

Because of the social nature of waterpipe smoking, sharing the waterpipe, a popular practice among youth worldwide can be associated with infectious disease risks, such as TB. Lastly, evidence suggests that waterpipe smoking can undermine tobacco control, as it can be used as a replacement for cigarettes among quitters, or serves as a gateway to cigarettes. A new study among 762 Danish youth (14–16 years) provides the first prospective evidence that waterpipe use predicts progression to regular cigarette smoking among Danish youth.

Although we still lack high quality studies of the health effects of waterpipe smoking, converging lines of evidence point to its deleterious health potential to smokers and non-smokers alike, as well as to its potential to undermine tobacco control measures.

Waterpipe's Addictive Potential

In 1997, Macaron and colleagues firstly demonstrated nicotine exposure in waterpipe users by measuring cotinine in their urine. Review of published studies of nicotine exposure associated with waterpipe use shows that daily waterpipe smoking produces a urinary cotinine levels equivalent to smoking 10 cigarettes per day. More recently, direct measurement of plasma nicotine in daily waterpipe smokers shows a five-time increase from

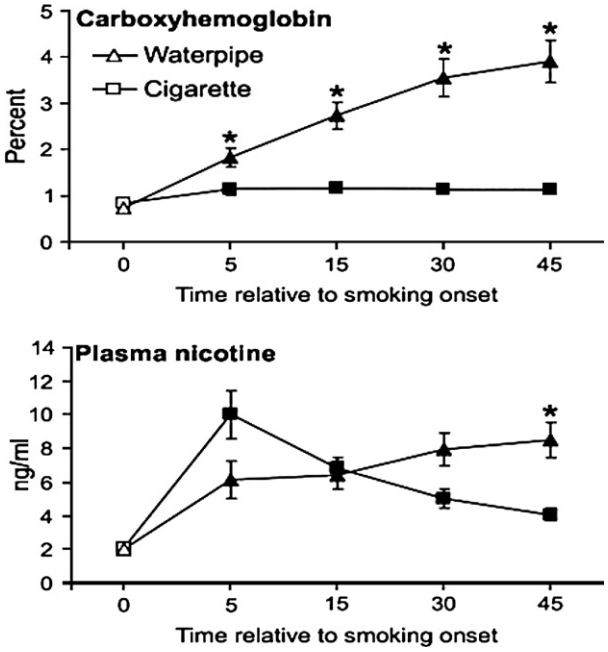


FIGURE 90.5 Mean carboxyhemoglobin (COHb) (top), plasma nicotine (bottom) of 31 subjects who smoked a WP (45 min) or cigarette (5 min). $^{*}P < 0.001$ between conditions. *Reproduced from (Eissenberg and Shihadeh 2009).*

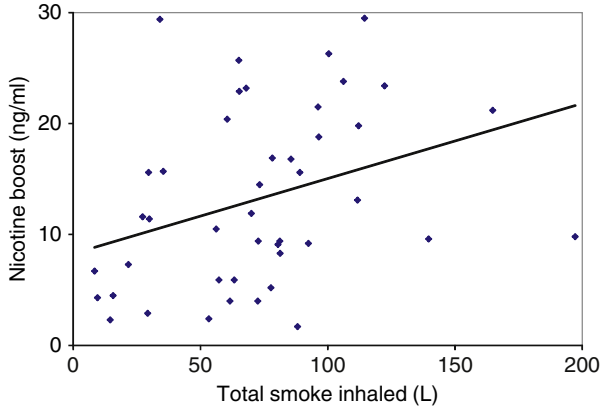


FIGURE 90.6 Correlation (Pearson $r = 0.34$) between plasma nicotine boost and total smoke inhaled (log transformed). *Reproduced from (Maziak et al., Addictive Behaviors, 2011).*

$3.07 \pm 3.05\ \text{ng ml}^{-1}$ pre-smoking to $15.7 \pm 8.7\ \text{ng ml}^{-1}$ post-smoking levels. Such waterpipe-associated nicotine exposure can exceed that associated with cigarette smoking. This was tested directly by comparing nicotine exposure of waterpipe and cigarette in dual cigarette/waterpipe smokers using crossover design. In this study conducted by Eissenberg and Shihadeh (2009), analysis of plasma nicotine area under the curve shows 1.7-time nicotine exposure dose for the waterpipe compared to cigarette conditions. Further support for the waterpipe as an important source for nicotine exposure comes from the demonstrated strong correlation ($r = 0.34$) between plasma nicotine levels and puff topography parameters such as total smoke inhaled during a waterpipe smoking session (Fig. 90.6).

Other than nicotine exposure, waterpipe smokers report known features of dependence, such as drug-seeking behavior, use escalation with time, self-perception of being hooked, and inability to quit despite repeated attempts. For example, a study conducted by Ward and colleagues (2005) in a random sample of 268 waterpipe users in Aleppo (Syria) showed that 28% wanted to quit and 59% reported an unsuccessful quit attempt in the past year. Belief in one's ability to quit was inversely related to perceived dependence. Another study from the same dataset conducted by Maziak and colleagues (2004) looked at waterpipe smoking frequency, as a surrogate marker of dependence, showed that dependent waterpipe smokers engage in behavioral adaptations to ensure access (e.g. carrying their own waterpipe with them) similar to what is seen in cigarette smokers. This early work at the Syrian Center for Tobacco Studies (SCTS) helped advance a bi-dimensional concept of waterpipe dependence; the first reflects the effects of nicotine, and the second reflects waterpipe's unique social dimension with more dependent waterpipe smokers increasingly showing solitary and home-based use patterns. Case histories from waterpipe users provide another line of evidence to the addictive nature of waterpipe; "It went from something fun I did each week, to each day, to 5–6 times a day. It became an addiction", or "I think I'm addicted to the social aspect of hookah".

As dependence becomes recognized among waterpipe smokers, studies looking specifically into this issue are emerging. For example, a Lebanese group aimed to develop and validate an 11-item Lebanon Waterpipe Dependence Scale, which included four domains; nicotine dependence, negative reinforcement, psychological craving, and positive reinforcement, and was correlated to waterpipe use frequency. Support of the dependence-inducing potential of waterpipe is beginning to gather from clinical lab evidence as well. In one study conducted at SCTS (2009), several subjective measures of withdrawal, craving, and nicotine effects were assessed pre-/post-smoking in 24 h-abstinent waterpipe smokers ($n = 61$, mean age 31 years). This study showed that that waterpipe smokers experience abstinence-related symptoms such as urges to smoke, restlessness, craving, and that these symptoms are suppressed significantly with subsequent waterpipe use. Such results, in addition to the anecdotal evidence of replacement potential between waterpipe and cigarettes (e.g. during cigarette quit attempts), prompted a further comparative study looking on how waterpipe, compared to cigarettes, suppresses craving and withdrawal symptoms. In this study, 12-h abstinent dual waterpipe/cigarette smokers

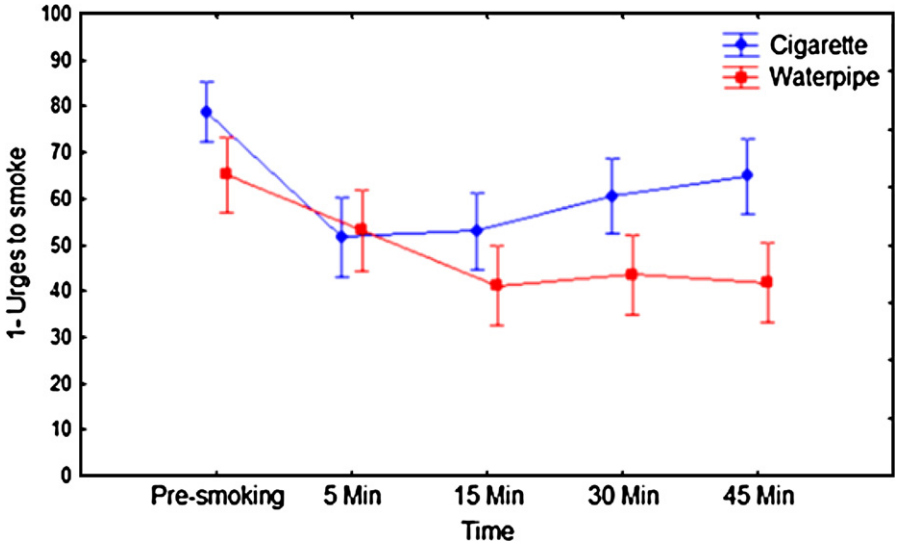


FIGURE 90.7 Mean scores of the item "Urges to smoke" in abstinent dual cigarette/waterpipe smokers ($p = 0.8$ for comparison at 5 min using repeated model ANOVA). *Reproduced from (Maziak et al., Addictive Behaviors, in press).*

(smoke waterpipe ≥ 1/week and cigarettes ≥ 10/day) smoked either waterpipe or cigarettes in two randomly-ordered sessions. Results show that waterpipe and cigarette smoking suppressed craving comparably at 5 min (end of cigarette smoking time point) (Fig. 90.7).

These data are consistent with the notion that waterpipe smoking is associated with features of dependence, and that waterpipe and cigarettes may be interchangeable in their effects on tobacco-abstinence symptoms. The clear public health implication of such information, is that waterpipe smoking is addictive and can hook users on nicotine, as well as thwart smoking prevention and cessation efforts among youth.

THE INTERVENTION AND POLICY IMPLICATIONS OF THE WATERPIPE EPIDEMIC

The Need to Understand Waterpipe Dependence among Youth

Despite what we already know about tobacco use, prevention and cessation interventions for youth have been hindered by the lack of clear understanding about how they get hooked on tobacco. Such understanding is needed to guide the timing, composition, and intensity of tobacco control interventions. While information about the initial development of tobacco dependence is beginning to emerge for cigarette smoking, knowledge to guide waterpipe interventions continues to be lacking. A Cochrane review by Maziak and colleagues failed to identify a single waterpipe cessation intervention, the development of which will require clear understanding of salient symptoms/domains of dependence in waterpipe users, and how they progress in their smoking habit.

An important aspect to consider in this regard is that the waterpipe has unique use patterns/features (e.g. intermittent use and prolonged sessions, preparation time, accessibility, and convivial experience of group use) that will likely shape dependence and how it is manifested in waterpipe smokers. For example, since the waterpipe is not as portable and accessible as cigarettes, waterpipe users may engage in more intensive behavioral adaptations to ensure access, such as carrying one's own waterpipe to places, and selecting cafés based on waterpipe availability. Other specific features such as waterpipe's sensory stimuli (e.g. smoke's aromatic smell) and social ambience can become cues for smoking, as well as a motivation for initiation. Moreover, the size of waterpipe requires deep inhalations in order to generate smoke and keep it going (puff volume $\approx$ 500 ml vs. puff volume $\approx$ 50 ml for cigarettes, Table 90.1). This feature combined with the less irritating smoke, compared to cigarettes, can lead to higher exposure to nicotine and toxicants during the experimentation stage and perhaps a faster path to dependence. As the waterpipe is becoming the first mode of contact with tobacco for many youth, the balance between dependence and access may determine which waterpipe users are likely to initiate cigarette smoking (i.e. the more dependant waterpipe smokers will likely not be able to abstain from smoking until they can smoke the next waterpipe and may switch to the more accessible cigarette).

Therefore, because of waterpipe's specific features and use patterns, initiation and dependence development in waterpipe smokers will likely follow

a distinctive path that need to be deciphered in order to develop effective interventions to curb waterpipe use.

Policy Aspects of the Waterpipe Epidemic

So far, data presented in this chapter not only indicate that waterpipe smoking has become a public health threat, but that this is perhaps the first tobacco use method since the cigarette that is showing all signs of a burgeoning global epidemic. They also indicate that the waterpipe is running its specific epidemiological course, and that its harmful and addictive profiles are likely to be shaped by its unique features and use patterns. Despite these alarming trends, most national and international tobacco control strategies (e.g. clean indoor air policies, prohibition of advertisement and sales to minors, large or graphic warning labels, taxation) still do not clearly address this tobacco use method. Furthermore, the waterpipe is not currently regulated (e.g. by the US Food and Drug Administration, or the Canadian Food Inspection Agency), and the content and packaging of waterpipe tobacco, or other accessories sold on the market, are not standardized. For example, many waterpipe users are driven by a misperception of reduced harm/addictiveness, which can be re-enforced by deceptive descriptors that appear on WP tobacco packages (e.g. "contains 0% tar and 0.05% nicotine") or accessories such as the charcoal (e.g. "smokeless and odorless", "free of chemicals", "100% natural"). All these descriptors aim to create an impression of a healthy product, and should be countered by appropriate packaging and advertisement policies.

The Framework Convention on Tobacco Control (FCTC), which represents the first global public health treaty, recognizes the diversity of tobacco products worldwide and calls for their inclusion in any tobacco control policy. However, most of the articles of the FCTC were formulated with the cigarette and the tobacco industry in mind. The picture becomes less clear when it comes to the waterpipe, where there is no well-defined industry behind it, and where the method itself is distinctive from cigarettes in many ways. Manufacturers and retailers of waterpipe's different parts/tobacco, and most importantly waterpipe cafés owners are mostly small businesses forming a multiplicity of converging interests rather than a well-defined industry or interest group. Given the current rate of waterpipe spread, we may still see the emergence of a more organized and powerful waterpipe industry and/or lobbying bodies on behalf of waterpipe interests.

While this is good news, because it means that, at least for now, waterpipe control efforts will unlikely face significant and organized opposition, understanding waterpipe's distinctive features – from cigarettes – is instrumental for the development of successful response to the global waterpipe epidemic. For example, banning traditional tobacco advertisement (billboards, print media) will likely have little influence on the waterpipe since much of its promotion and sales are done over the Internet and social networks, or implicitly as part of cafés/restaurants promotion. In addition, most of warning label/pictorial policies make no sense for the social waterpipe smoker because in a café setting the waterpipe is usually served prepacked (i.e. the user is not exposed to any warning put on the waterpipe tobacco package). In the same café setting, unclear policies concerning minors' access to the waterpipe may allow underage people to order waterpipe in the café as a way of "sampling" rather than overt purchase of waterpipe tobacco. The call for cessation services by the FCTC furthermore, requires evidence-based cessation programs for the waterpipe, which do not exist currently. This will not only influence waterpipe control efforts, but will likely reflect on cigarettes' as well. Unless the waterpipe is addressed adequately in smoking cessation programs, evidence shows that cigarette quitters may turn to the waterpipe to deal with unpleasant abstinence symptoms. Finally, in some parts of the world (e.g. United States), waterpipe venues can be benefiting from tobacco control policies (e.g. smoking bans in restaurants and bars), because they can be exempted under certain provision as a "retail tobacco establishment".

As mentioned earlier, much of the promotion of waterpipe is exploiting the harm-reduced notion contingent on the assumed filtering function of water. In an attempt perhaps to circumvent potential waterpipe tobacco control policies, this concept has been taken further recently with the introduction of non-tobacco-based waterpipe herbal products (e.g. Soex, a sugar cane-based product). These have been marketed with claims of "same flavorful smoke found in other shisha without the harmful effects of tobacco". Comparative analysis of this product with regular waterpipe *Maassel* using a crossover design shows that Soex delivers a similar amount of CO but not the nicotine, making it a bad "harm-reduced" option, and supports the need to regulate these products against deceptive marketing strategies (Fig. 90.8). Exploiting the harm-reduced hype is not confined to waterpipe tobacco, but is involving the artifact itself as well in a way reminiscent of the electronic cigarette. An electronic variant of the waterpipe has appeared recently on the market (called Halooka™) with claims of being healthier than the original waterpipe through the application of electronic heating source and nicotine containing cartridges (variety of flavors of course). Apparently, because the electronic waterpipe does not involve ignited tobacco, it will fall outside most of tobacco control regulations,

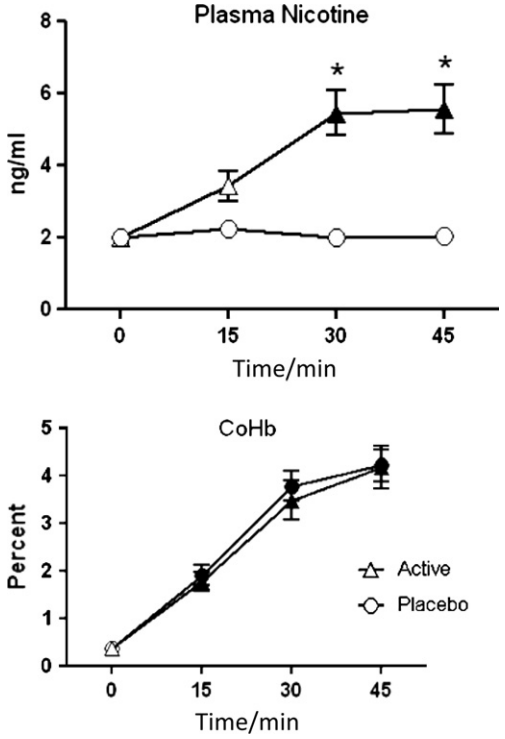


FIGURE 90.8 Mean ± SEM for plasma nicotine (top), COHb (bottom) for active and placebo waterpipe tobacco conditions in 37 waterpipe smokers who completed two randomly-ordered waterpipe sessions using either regular tobacco or placebo (tobacco free herbal waterpipe preparation). Filled symbols indicate a significant difference from baseline and asterisks (*) indicate a significant difference between active and placebo. *Reproduced from (Eissenberg et al., SRNT, 2010).*

regardless of what are its potential harms or its role in sustaining the smoking culture and nicotine addiction amongst youth.

Accordingly, policy initiatives such as health warnings, bans on advertisement/deceptive descriptors, cessation, price/taxation, and controlling minors' access should accommodate the multi-component, electronic media driven, and the social nature of the waterpipe. It should also accommodate the fast changing landscape of waterpipe products and promotion that exploit the social networks media and harm-reduced perception to thwart potential waterpipe tobacco control efforts.

CONCLUSIONS

Given waterpipe's global reach, its potential to hook young people on nicotine, to replace cigarettes among quitters, and to harm smokers and non-smokers alike, the threat of waterpipe smoking to global public health need not be overemphasized. Understanding waterpipe's unique features and how these influence users as they progress in their smoking habit, as well as the dynamic environment that supports waterpipe use and promotion, will be essential for the development of prevention, treatment, and policy interventions to curb the global waterpipe epidemic.

SEE ALSO

Maziak, W., 2011. The global epidemic of waterpipe smoking. Addictive Behaviors, 36(1–2), 1–5.

List of Abbreviations

CO	carbon monoxide
COHb	carboxyhemoglobin
COPD	chronic obstructive pulmonary disease
EM	Eastern Mediterranean
FCTC	Framework Convention on Tobacco Control
FEF 25–75%	forced expiratory flow 25–75%
GYTS	Global Youth Tobacco Survey
NO	nitric oxide
PAH	polycyclic aromatic hydrocarbons
PEFR	peak expiratory flow rate
SCTS	Syrian Center for Tobacco Studies
SES	socioeconomic status

Glossary

Waterpipe is a centuries-old tobacco use method with an ambiguous origin, but is often associated with Middle Eastern societies. Although known by different local names (e.g. hookah, narghile, shisha), the term waterpipe has been coined in order to combine tobacco use methods that share the unifying feature of passage of tobacco smoke through water before inhalation by the smoker.

Maassel is a wet mixture of tobacco, sweeteners (e.g. molasses), and flavorings that is currently most commonly used in the waterpipe.

Further Reading

Akl, E.A., Gaddam, S., Gunukula, S.K., Honeine, R., Jaoude, P.A., Irani, J., 2010. The effects of waterpipe tobacco smoking on health outcomes: a systematic review. International Journal of Epidemiology 39 (3), 834–857.

American Lung Association, 2009. Tobacco policy trend alert an emerging deadly trend: waterpipe tobacco. www.slati.lungusa.org/alerts/Trend%20Alert_Waterpipes.pdf.

Cobb, C., Ward, K., Maziak, W., Shihadeh, A., Eissenberg, T., 2010. Waterpipe tobacco smoking: an emerging health crisis in the United States. American Journal of Health Behavior 34 (3), 275–285.

Doubeni, C.A., Reed, G., Difranza, J.R., 2010. Early course of nicotine dependence in adolescent smokers. Pediatrics 125 (6), 127–133.

Eissenberg, T., Shihadeh, A., 2009. Waterpipe tobacco and cigarette smoking: direct comparison of toxicant exposure. American Journal of Preventive Medicine 37 (6), 518–523.

Khabour, O.F., Alsatari, E.S., Azab, M., Alzoubi, K.H., Sadiq, M.F., 2010. Assessment of genotoxicity of waterpipe and cigarette smoking in lymphocytes using the sister-chromatid exchange assay: a comparative study. Environmental and Molecular Mutagenesis [Epub ahead of print].

Knishknowy, B., Amitai, Y., 2005. Water-pipe (narghile) smoking: an emerging health risk behavior. Pediatrics 116 (1), e113–e119.

Martinasek, M.P., McDermott, R.J., Martini, L., 2011. Waterpipe (hookah) tobacco smoking among youth. Current Problems in Pediatric and Adolescent Health Care 41 (2), 34–57.
Maziak, W., Ward, K.D., Eissenberg, T., 2007. Interventions for waterpipe smoking cessation. Cochrane Database of Systematic Reviews (4), CD005549.
Maziak, W., Eissenberg, T., Ward, K.D., 2005. Patterns of waterpipe use and dependence: implications for intervention development. Pharmacology, Biochemistry and Behavior 80 (1), 173–179.
Maziak, W., Ward, K.D., Afifi Soweid, R.A., Eissenberg, T., 2004. Tobacco smoking using a waterpipe: a re-emerging strain in a global epidemic. Tobacco Control 13 (4), 327–333. http://dx.doi.org/10.1136/tc.2004.008169.
Maziak, W., 2008. The waterpipe: time for action. Addiction (11), 1763–1767.
Maziak, W., Rastam, S., Ibrahim, I., Ward, K., Shihadeh, A., Eissenberg, T., 2009. CO exposure, puff topography, and subjective effects in waterpipe tobacco smokers. Nicotine & Tobacco Research 11 (7), 806–811.
Maziak, W., 2011. The global epidemic of waterpipe smoking. Addictive Behaviors 36 (1–2), 1–5.
Nakkash, R., Khalil, J., 2010. Health warning labelling practices on narghile (shisha, hookah) waterpipe tobacco products and related accessories. Tobacco Control 19 (3), 235–239.
Neergaard, J., Singh, P., Job, J., Montgomery, S., 2007. Waterpipe smoking and nicotine exposure: a review of the current evidence. Nicotine & Tobacco Research 9 (10), 987–994.
Raad, D., Gaddam, S., Schunemann, H.J., Irani, J., Abou Jaoude, P., Honeine, R., Akl, E.A., 2010. Effects of waterpipe tobacco smoking on lung function: a systematic review and meta-analysis. Chest [Epub ahead of print].
Shihadeh, A., 2003. Investigation of mainstream smoke aerosol of the argileh water pipe. Food and Chemical Toxicology 41 (1), 143–152.
Shihadeh, A., Saleh, R., 2005. Polycyclic aromatic hydrocarbons, carbon monoxide, "tar", and nicotine in the mainstream smoke aerosol of the narghile water pipe. Food and Chemical Toxicology 43 (5), 655–661.
Ward, K., Hammal, F., VanderWeg, M., Eissenberg, T., Asfar, T., Rastam, S., et al., 2005. Are waterpipe users interested in quitting? Nicotine & Tobacco Research 7, 149–156.
Warren, C.W., Jones, N.R., Eriksen, M.P., Asma, S., Global Tobacco Surveillance System (GTSS) collaborative group, 2006. Patterns of global tobacco use in young people and implications for future chronic disease burden in adults. Lancet 367, 749–753.
Warren, C.W., Lea, V., Lee, J., Jones, N.R., Asma, S., McKenna, M., 2009. Change in tobacco use among 13–15 year olds between 1999 and 2008: findings from the Global Youth Tobacco Survey. Global Health Promotion 16 (Suppl. 2), 38–90.
World Health Organization (WHO), 2005. Waterpipe Tobacco Smoking: Health Effects, Research Needs and Recommended Actions by Regulators. WHO, Geneva, Switzerland.

Relevant Websites

Syrian Center for Tobacco Studies. www.scts-sy.org.
World Health Organization. www.who.int/tobacco/en/.

CHAPTER

91

The Cell Phone in the Twenty-First Century: A Risk for Addiction or a Necessary Tool?

Xavier Carbonell, Ursula Oberst, Marta Beranuy

FPCEE Blanquerna, Universidad Ramon Llull, Barcelona, Spain

OUTLINE

Social progress parallels that of human communication. The need to communicate led to the production of written languages, a plethora of communication channels, and has been a crucial element in the development of our brains. This chapter deals with the importance of one of these communication channels, the *cellular telephone*, and how it impregnates our culture. The cell phone is a portable device which has become a social object that is personal, exclusive, and intimate. The different names used for this device help us see the very particular relationship each country has with this tool: it is called a *móvil* in Spain, *celular* in Latin America, *cell phone* in the United States, *handy* in Germany, "*sho ji*" (handheld) in China, and "*makhmul*" (portable) in Arab-speaking countries. In Japanese society, this technology is so commonplace that cell phones are simply called telephones. Whatever its name, people everywhere have developed an intense relationship with their cell phone, much more intense than they ever did, in its day, with an ordinary (fixed-line) telephone.

The fusion of computing and telecommunications in the 1970s gave rise to the development of the so-called information and communication technologies (ICT). This moment saw the birth of the philosophy of mobile telephony, a technology which combines the Bell telephone, Morse telegraph, Marconi radio, and computing. Humans have always tried to overcome distance as

Principles of Addiction, Volume 1
http://dx.doi.org/10.1016/B978-0-12-398336-7.00091-7

a barrier to communication, giving rise to carrier pigeons, telegrams, letters, postcards, books, magazines, radio, television, fixed-line telephones, electronic mail, Internet Relay Chats (IRC), and videoconferencing. The latter few overcome more than just the distance barrier, achieving real-time communication at a distance.

Cell phone use in today's society is common, and sales have increased notably. The number of cell phone subscribers has risen all over the world, Europe being the continent with most lines (followed by Oceania, America, Asia, and Africa). Between 2002 and 2003 the total cell phone users worldwide crossed the 1000 million boundary. It had taken 130 years for fixed-line–based telephony to reach this figure, whereas cell phones did it in a decade. Eight of every ten people in Europe possess a cell phone. Many people who only a few years ago would not buy one, or used one only sporadically, nowadays use one daily. Some models are truly objects of desire and have provoked changes in forms of communication. Cell phone uses even include things like teaching English, helping people give up smoking, or therapy in people with driving phobia or agoraphobia. Cell phones have played an important role in communication between opposers of the totalitarian regimes of northern Africa in 2011 and in the diffusion of images of their repression.

In parallel with increases in use and social acceptance, there has also been consideration of the possible negative consequences: medical (negative effects of the electromagnetic waves emitted by cell phones themselves and by their base station antennas), in road safety (effects and risks of cell phone use while driving or cycling and their involvement in traffic accidents), and psychological. Cell phones can play an important role in bullying (recording and diffusion of acts of vexation), in violence (recording and diffusion), and in sexual harassment. One of the most worrying negative consequences is the possibility of creating addiction, particularly among teenagers and young people. This possibility will be dealt with in the next section.

WHAT MAKES CELL PHONE USE SO GRATIFYING?

One of the factors determining the capacity of a substance to create addiction is its gratifying properties. It is accepted that the more intense the positive reinforcement, and the shorter the delay between consumption and physiological response, the greater is the capacity of a substance to produce addiction. So, let us analyze the positive reinforcement properties of cell phone use.

Euphoria

Euphoria from cell phone use appears to be at least as strongly related to message emission (calling or sending someone a message) as reception is to being feeling valued or loved when calls or messages are received.

Instrument

A cell phone may at the same time serve as pocket watch, alarm clock, digital camera, sound and/or video recorder, electronic diary, video console, radio, mp3 player, or Global Positioning System (GPS). It is a multifunction instrument with many utilities adapted to the age and social role of its owner.

Symbol of Identity

Cell phones have become one more element among the intimate components which constitute the personal sphere (just as other things do, such as a wristwatch, wallet, photos, key ring, etc.) with which the bearer has an emotional bond. Never before has a technological apparatus come to have such importance in so many people's daily lives, so essential to revealing identity.

The degree of personalization possible with a cell phone is one of the factors favoring expression of individual identity, particularly among younger people. The cell phone appears to have become an object through which a person can provide clues about their gender identity, social and professional position, attitude toward society, character, personality, or mood. A cell phone, like clothes, may transmit information about an individual's characteristics and about the idea they have of themselves, and which they want to transmit to others.

Social Status

Cell phone technology appears to confer power on young people. A young consumer who buys a cell phone feels powerful, not only through the use made of it, but also through the purchase in itself. Also, the number and/or quality of messages received, the number of calls, the number of contacts in the address book, the sophistication of the games and services offered by the cell phone, and the brand of the apparatus, all help to enhance the user's social status.

Social Network

Cell phones also are a tool to build a social network via the device's contact list. These networks are constantly evolving; the rate at which new phone numbers are added to contact lists is as rapid as that

of numbers falling into disuse. Moreover, we may speak of a collective identity. The social networks based on cell phones have created a new sense of identity for teenagers and young people.

Online Social Networks

In its short life, the cell phone industry has managed to adapt to keep pace with the demands of users, and create new needs: cell phones for professionals and business people, cell phones with only basic functions for children and the elderly, others for listening to digital music, and finally the cell phone as a discrete medium for checking emails and *online social networks*, and consulting Internet.

Independence

Cell phones play a significant role in socialization and creating a feeling of belonging to a group, particularly among teenagers. They foster a process of emancipation from parents and act as a kind of barrier between teenagers and their parents. In other words, for young people the cell phone is above all a personal telephone, and having a personal telephone not accessible by parents marks a boundary. Having a cell phone helps a teenager to acquire an ever greater sense of self and an increasing orientation toward the peer group. The cell phone favors independence and reinforces contact with friends and other people outside the family.

The cell phone is definitive when it comes to deciding if a young person can enter into society so that young people use them to maintain their social framework. At the same time the social contacts allow them to maintain their status in terms of class and peer group. In Tokyo public schools cell phone owners have more friends than nonowners. Children and teenagers usually receive a cell phone as a gift from parents. Making a present of a cell phone could be seen as a rite of passage, a gift related to initiation into the teenager phase, into social independence. And it would appear that this rite occurs at ever earlier ages.

Short Distance

The cell phone is an instrument which facilitates contact over short distances, in the sense of contact with people with whom we do not relate on a daily basis. One typical characteristic of youth is that it involves tightly closed, local social circles, neighborhood, school, club, etc., and the cell phone here is a practical medium for maintaining contact when face-to-face conversation is not possible. When members of the social network are a greater distance away, other communication channels are preferred, such as electronic mail, social networking, fixed-line telephone, or, ever more rarely, letters.

Increased Security and Control

Cell phones are instruments of control which generate feelings of security among parents, between couples, or even for oneself when away traveling. Parents buy their children cell phones because of a need to control them and restrain them with a *"digital leash."* This eagerness to control is a feeling which parents transmit to their children. Often the degree of control and sense of security is false: it is very easy to lie about where one is, and, anyway, battery life and coverage are both limited. Teenagers can mutually communicate while at home without parental control, something which was more difficult with ordinary (fixed-line) telephones, as well as at school without being controlled by teachers. In adults, this sensation of controlling/being controlled also occurs in sentimental and workplace relationships.

Permanent Mobility and Access

The fact of being mobile means that accessibility of people carrying a cell phone is perceived as permanent. This process gives rise to two opposing illusions: one, believing that we are not being controlled when in fact we can be located at virtually any moment wherever we are, and two, believing that one can control others when really cell phones only allow us to hear someone's voice or receive an SMS without us really knowing from where. Even so, parents prefer to believe they have some degree of control than to let their children escape from their clutches. Something similar could occur in sentimental and workplace relationships between adults, and some firms employ cell phones with GPS functionality permitting their location to be tracked.

When the user does not answer calls or respond immediately to SMS text messages, the caller can experience a sensation of concern. This arises because of the wrong interpretation of availability, often understood as "*obligatory*" creating an illusion of "*permanent availability*" and the cell phone user is pressured to carry the device turned on, and to always respond to calls or SMSs. Even so, the user is more interested in being able to call others while on the move, lending less importance to always being locatable.

Entertainment and Games

Cell phones carry a broad range of functionalities, and may even act as a portable videogame console.

Being up-to-date, playing the latest games, feeling integrated, and/or up with the fashion are goals pursued by many of today's teenagers and young people. One must also bear in mind that, increasingly, more leisure time activities are available through cell phones, for example betting, buying, getting sexually stimulated, and downloading music and videos. Children under 10 years of age regard games as the most important characteristic of cell phones, since at their age communication in itself is too abstract. The incorporation of applications ("*Apps*") in the latest generation of cell phones (so-called *smartphones*) has opened up an enormous range of possibilities for their use at work, for leisure, and for practical aspects of daily living; in many of these applications, these functions are intermixed. Cell phones are becoming personal mobile computers.

Synchronous and Asynchronous Communication

Voice calls and text messages are used differently depending on the purpose and on the characteristics of the message sender and receiver. Voice is synchronous communication, simultaneous in time, whereas text messages are asynchronous, like electronic mail.

Family Conciliation

The social evolution of family structures could partly explain the increase in personal telephone use. We may speak of various factors: (1) the emergence of single-parent or patchwork families: particularly reliant on external phone connections due to characteristics of their structure; (2), internal democratization of the family which accentuates individual autonomy and is susceptible of favoring diffusion of a less collective, more personal, form of telephonic communications; (3) the demand for individual communications devices given that children remain living with their parents for longer in some countries. The increasing incorporation of women into the workforce might be considered a fourth factor. Even though cell phones have not changed any social conventions, women tend to use it to cope with family responsibilities across a space-time gap, bringing their private world of domestic responsibilities to their public, occupation-related world, and vice versa.

All these changes in family morphology are reflected in affective and social bonds. Cell phones have made it possible for teenagers to construct a kind of *virtual brotherhood.* Moreover, cell phones promote individual thinking and networks of external support and propitiate virtual proximity (in the double sense of the word virtual). Connections mediated by cell phones only deal with the issue generating the call, leaving the parties involved free of any emotional commitment beyond the topic dealt with in the conversation or message. Present society demands fast and efficient connections. In this respect, distance is not an obstacle for connecting, but being connected is likewise not an obstacle for maintaining distances. Being connected is more economical than really relating. Thus, we may speak of new family constellations and emotional processes deriving in a society, still under construction, which gives rise to new ways of communicating to maintain family unity and the sense of belonging which both adolescents and adults need.

Individualization of Assets

This is one aspect of social evolution and the increasing quality of life in the Western world. In the technological field, telephones have followed the same path as television, in becoming an individual asset rather than a family one. Just as teenagers may have a television set in their own bedroom, they also have their own computer, and cell phone, etc.

WHAT IS THE SOCIAL SIGNIFICANCE OF TEXT MESSAGING?

The short message service (SMS) is the facility which permits sending text messages between fixed-line or mobile phone devices. The SMS culture has hit our society and their use is a social phenomenon: inviting users to political meetings or parties, follow-up of patients, as vehicle for therapy, obtaining status of commercial and bank transactions, participating in television contests and programs, receiving official bulletins, etc. Let's consider the inherent characteristics of the SMS separately.

Functionality of Written Language

Written information has an added value: its permanence. SMSs, with this attribute, are thus very different from oral conversation. The transmission of information, petitions, and self-expression represent the genuine, and historically almost invariable, dimensions of communication. An SMS is not a letter, but could be considered the equivalent of the postcard or telegram at least in regard to brevity and condensed content, with the additional advantage that they may be sent to various people at a time. An SMS can be a substitute for electronic or postal mail. At Christmas and New Year people use

skyrockets, as they are a quick and instantaneous replacement for Christmas cards.

Expression of Feelings

SMSs help to express, with little direct involvement, the feelings that people cannot or do not want to express orally and moreover they respond to the impulsive need to share feelings at the moment they are being felt. Also, regardless of age, sending an SMS indirectly implies that one is manifesting their very presence to the receiver, and thus the SMS carries an important symbolic load. This all means that SMSs are perceived as particularly satisfactory and, in the view of some authors, end up promoting more intimate ties and enriching personal relationships.

Abbreviated Language

The need to communicate as much as possible in the reduced space of an SMS has contributed to the development of an intensive use of abbreviations. In this form of expression, all kinds of strategies are used to abbreviate as much as possible, for example "lol" to mean "laughing out loud" or "xD" to mean happiness or laughing very hard. Lists giving the translation of these abbreviations abound, as do certain rules for writing SMSs.

Use of Emotion Icons

Diagrammatic representations of emotional states, (for example a smiling face, whether written by using ordinary keys :-) or via the ☻ symbol, may be used to indicate happiness. These icons, known as *emoticons*, serve to express feelings in the middle of the written text. The advantage of emoticons is more notable in an SMS than in electronic mail because the available space is more limited. How better to save words than to replace them with a pictorial representation.

Nocturnal Networks

SMSs can be used to set up a virtual nocturnal network of friends. Whereas most adults use SMSs mainly to confirm appointments, teenagers use them to express a broad spectrum of emotions and feelings which result when they find themselves alone, usually just before going to bed. With the responses to their SMSs, also charged with romanticism, they feel that their emotions have been corresponded, and hence they feel valued. As a result of all this, there is a tendency for young people to save emotion-charged messages in the cell phone's memory.

Avoidance of Telephone Conversations

The unilateral aspect of an SMS, and its concise, direct, and synthetic nature responds simultaneously to three needs: to save time, save money, and, most interesting of all, maintain bonds even when the user does not want to get into a telephone conversation, due to the degree of commitment a voice call involves.

Respect for Privacy

The beep notifying reception of an SMS is usually shorter and more discreet than that of a normal call. SMSs are an easy form of communication for shy people, or people in embarrassing situations. Some people politely send an SMS before calling, to check that the other person is available and wants to speak with them.

Anxiety

Some people feel uncomfortable or irritation when they do not get a response to an SMS they have sent. This could be due to a variety of factors: the immediateness and permanent availability, the particularity of written language, the exclusive dedication needed to send an SMS (typing an SMS requires dedicating time exclusively to its composition, and hence of thinking about the person for whom it is intended, whereas making a call permits doing several other things at the same time). A user faced with an unanswered SMS could feel the time spent writing it has not been corresponded and interpret that the investment in involvement has likewise not been corresponded by a response of similar intensity. This can lead to increased anxiety. To cope, some users use the missed calls technique with the aim of attracting the receiver's attention so that they realize they have received an SMS, or in order to make them understand the need for an immediate reply. This can lead to the creation of a loop, and can thus escalate levels of concern until a state of genuine anxiety is reached. Some users even go further, in a desperate attempt to get a reply from anyone at all, by sending an SMS to an entire list of contacts, in this way, as might be expected, increasing their anxiety even more.

DIFFERENCES IN USAGE IN TERMS OF GENDER AND AGE

Possession of a cell phone nowadays is independent of age group and gender, although their preferred modes of use differ. In regard to gender, cell phone use by girls is characterized by being mainly to keep up with their social network, whereas boys use it more to

coordinate their movements, and to play games. The structure of social relationships mediated by cell phones coincides with the typical differential gender characteristics: (1) women extend their social networks through the use of SMSs more than men; (2) women use cell phones to maintain social contacts whereas men use them for commercial gain, similar to what already happens with fixed-line telephone use; and (3) women are more compulsive in their cell phone use than men. Among teenagers, for example, girls' cell phones are used as a security measure and for controlling their autonomy, whereas among boys cell phone use is related to a process of independence and gender identity charged with the symbolism of modernity. In some countries, women use cell phones more than men precisely because they have less freedom of movement.

Young, highly educated adults also face a particular form of socialization. They aspire to company leadership positions and cell phones are well suited to bolstering their image, epitomized by giving and receiving orders while running between flights. In fact, business people tend to use cell phones more when on the move than in other circumstances. The issue here is about marking the difference between oneself and the people around us, and parallels the teenager drawing a boundary line between themselves and their parents. Adults tend to prefer voice calls due to their synchronicity and because they are simple and practical. Nor is there so much personalization of the device, perhaps because they correspond to a different generation whose need for establishing an identity is a phase that has been passed. Still, adults are not entirely strangers to fashion or to using a cell phone as a symbol of status and identity. For example, think of the migration from company-provided *Blackberrys* toward *iPhones*, from the micro-keyboard to touch screens, etc.

IS THERE SUCH A THING AS ADDICTION TO CELL PHONES?

To date there is little scientific literature with reliable data on prevalence, symptoms or clinical cases of addiction to mobile telephones, yet there is a climate of social alarm, generated by mass media insistence on their addictive risks.

The colloquial usage of the word addiction can be confused with its technical usage. One initial possibility is that the supposed addiction to cell phones is a problem which is limited in time and in severity of its consequences. It may arise through a "novelty effect," where stimulation by something novel provokes an increase in the frequency of a behavior during a short period of time, after which the behavior becomes less frequent or disappears. Something similar to what can happen, say, with the purchase of a new camera, or bicycle, or attendance to a fitness or wellness center. A period of adaptation to the new technology is also necessary, not only for the user but also their immediate social circle (in many cases the parents of the user). A second possibility which likewise must not be confused with true addiction is the non-severity of the consequences. Abuse of a cell phone can generate discipline problems at school (paying attention to the device when not allowed) or problems with parents if the bill is too high. However, these problems cannot be considered equivalent to those caused by an addiction to substances and must be seen as comparable to other limits which teenagers need to have imposed on them as part of their maturing process. Just as the behavior of biting fingernails is a bad habit but not an addiction, and although it may be considered a problem of control of impulses, in fact it is not treated as such in the classification of mental disorders.

Another aspect is that this type of pathological cell phone use would only be possible in people suffering a mental disorder or primary personality disorder. Pathological cell phone use would be a symptom of depression or an impulse control disorder, and would not be observed in healthy people. A third aspect would be, as with Internet, not to confuse addiction *to* a device with an addiction *on* a device. A pathological gambler who uses Internet or a cell phone to place bets is addicted to gambling, not to Internet or cell phones. Similarly, addiction to phone sex must not be confused with addiction to telephones. Finally, some cell phone users confuse dependence on a technology with a symptom of addiction. For example we cannot do without servo-assisted steering in our car, or electricity in our home or workplace. Similarly, a young person who feels the need to always carry their cell phone so as to be able to give warning in some emergency or simply to be available for receiving calls does not suffer true addiction, rather they are making conscious use of a security measure, just like a safety belt, or a car's servo-assisted steering or braking system.

When comparing problematic cell phone use to the well-known symptoms of addiction, among the most commonly described aspects, we may observe:

- Tolerance. The well-being originated by gratifying stimuli, such as receiving a call or SMS, is short-lived and reinforcing behaviors are repeated more often, such as calling insistently for no precise purpose, or soliciting further SMSs.
- Abstinence. As soon as the possibility to use the cell phone is lost, symptoms similar to a withdrawal syndrome appear. For example, a flat battery or a loss of coverage leads to displays of anxiety, general malaise, anger or uneasiness, and the same

may be felt when one does not receive a reply to calls made or messages sent.
- Insecurity. Some people are afraid of going out of the house without their cell phone, and would go back for it if they forget it; they may feel nervous or experience debilitating insecurity, and not be able to do anything when without their cell phone; children are particularly sensitive to developing uneasiness, even anxiousness, if they are obliged to do without their cell phone, whether this be as a result of a breakdown or as a punishment imposed by parents. This is related with the fear of losing something important, of being left out of the information circuits (fear of missing out, FOMO) or of missing that hoped-for or anxiously awaited call or SMS. Moreover, the user knows that other users expect one to always have their cell phone with them and fears that the others may be disconcerted if their calls are not answered.
- Attempts to control or cut usage. Some users attempt to control their cell phone use by blocking calls or setting quotas, by disconnecting the ring tone or switching the device off.
- Persistence in using the cell phone despite its negative effects. The most common of these are (1) spending more than initially intended (children can even fool, lie to, or steal from their parents); (2) using the cell phone in places where it is prohibited, or while driving; (3) use the cell phone so much it reduces time available for sleeping; and (4) have discipline problems in class or at school.

In any case, it appears that the symptoms found in the literature regarding pathological cell phone use are less consistent and less serious than those relating to Internet use. It is rare to find clinical cases of cell phone addiction. But it seems that there are certain maladaptive behaviors (or problematic uses) with respect to this medium. The reported prevalence rates of problematic cell phone use in population surveys vary from 2.8 to 10.4%. This problematic use was greatest in the youngest age groups. The results suggest that females have more difficulties with phone use than males and perceive their use as more problematic.

Some authors have conducted research into possible addiction to *instant messaging* (*IM*) among teenagers, i.e. one aspect of cell phone use. In a sample of 330 Chinese teenagers, 9.8% of them were classified as IM addicts; factor analysis identified four major addiction symptoms: preoccupation with IM, loss of relationships due to overuse, loss of control, and escaping from reality.

The distinction between information use, communication use, and identity-altered communication use could explain why cell phone use is not itself an addiction. The traditional use of cell phones has been for communication. Since calls and messages are exchanged with people whose identity is known, there is no identity-altered communication and therefore the risk of problematic and/or addictive use is likely to be very low. In identity-altered communication, playing with one's identity can become problematic and/or pathological as the users take on alternative (i.e. false) identities that provide greater satisfaction than their true self, allowing them to escape from their true self. In the case of cell phone use or Internet chat applications such as *Messenger*, the negative consequence is time wasted, while the positive aspect is maintenance of social relations with friends and acquaintances and broadening of the social network. However, this risk could potentially be higher for newer generation cell phones since applications that promote alteration of user identity may be supported.

For instance, some people may confuse or self-define dependence on a particular technology as an addictive behavior. For this reason, some people consider themselves cell phone addicts because they never go out of the house without one, do not turn it off at night, are always expecting calls from family members or friends, and/or they overutilize it in their work and/or social life. Finally, there is also the importance of economic and/or life costs. The crucial difference between certain forms of game playing and pathological game playing is that some applications involve a financial cost. If a person is using the application more and is spending more money, there may be negative consequences as a result of not being able to afford the activity (e.g. negative economic, job-related, and/or family consequences). High expenditure may also be indicative of cell phone addiction but the phone bills of teenagers are often paid by parents, therefore the financial problems may not impact on the users themselves.

The latest generation of cell phones, with permanent Internet connection, *web 2.0*, and a growing multitude of related functions (*Apps*, real-time emailing, *VoIP*, etc.), could increase the risks of problematic use, since they combine the elements of ICT and remove the clear demarcation between information and communication. Every day more and more people use their cell phones compulsively to check their email or SMS inboxes and social networks, play games, listen to music, or idly scroll the appstore in search for interesting and useful applications. whenever not occupied with something else, or indeed, even when they are. Thus the cell phone is becoming a catalyst of FOMO, and hence increase stress and anxiety.

Studies show that teenagers are the population most at risk of suffering the negative effects of cell phone use, and may need psychiatric help to avoid relationship and academic after-effects. It would also be beneficial to develop school-based preventive programs, aimed at

both students and their parents. For the former, through their tutors, promote well-adapted use of this communication medium. The way to tackle this would have to promote self-esteem, autonomy, self-concept, etc. in such a way as to ensure they have a favorable effect on the overall health of the adolescent. For parents, through seminars, facilitate guidelines for actions and support to follow-up, with the same objective.

SUMMARY

In conclusion, while perhaps it is not wise to label these possible problematic uses as addiction, one may assert that there is a series of maladaptive behaviors in regard to cell phone use, which change people's daily lives and play an important role in their feelings of personal security, identity, and belonging to a social group. While not intending to pathologize this behavior – we would rather classify it as problematic or maladaptive – the possible negative consequences in the long run must still be recognized. We suggest studying the possible negative consequences of mobile phone use not exclusively in the context of addiction, but in a broader context of the negative consequences for psychological health in a postmodern society, for three reasons.

First, we live in a society where, encouraged by the style of advertising, type of leisure activities and social values, high-impact, but short-lived emotions seem to be more valued than deeply felt and long-lasting sentiments. The new technologies and especially cell phones with texting are perfect channels for expressing these types of emotions in a quick and volatile ("light") way. They are more addictive than deeper feelings, for the same reason that gambling is addictive: its quick, but not always contingent recompense. Second, the overall possibility of permanent and global access to information creates in many people a feeling of "infoxication," an information overload impossible for the individual to cope with; but the fear of missing important details for their personal or professional lives keeps people in a constant state of concern about catching up that may ultimately lead to anxiety and stress.

And finally, the changing ways of relating to other people, identified as "liquid bonds" by the sociologist Zygmunt Bauman, not only allow, but also encourage the individual to create and dissolve social relationships easily, a life style that can lead to considerable psychological distress. Calling and texting behavior could become excessive precisely because it corresponds to contemporary communication styles and habits. In this sense, "addiction to the cell phone" can be understood as a social over-adaptation to the predominant values of our society in order to avoid being excluded from social dynamics: being always informed, being always available, but preserving continued possibility to avoid the other, to refrain from implication, and to elude compromise.

SEE ALSO

Internet: Immersive Virtual Worlds, Overuse of Social Networking, Video Game Addiction, Substance Use and Mental Health Issues on the College Campus, Historical Understandings of Addiction

List of Abbreviations

FOMO	fear of missing out
GPS	global positioning system
ICT	information and communication technologies
IM	instant messaging
IRC	Internet relay chat
SMS	short message service
VOIP	voice over Internet protocol

Glossary

Apps in general, short for "application software," a program designed for end users. The term was broadly introduced by iPhone™; app, or application, is what Apple® calls third-party software programs developed specifically for the iPhone™ and the iPodTouch®. The applications available can be downloaded directly by the cell phone, or downloaded to a computer and transferred to the phone. Other providers now also offer application software for their mobile phones.

Emoticon portmanteau of *emotion* and *icon*, a facial expression pictorially represented by punctuation and letters, usually to express a writer's mood. Emoticons are often used to alert a responder to the tenor or temper of a statement, and can change and improve interpretation of plain text.

Information and communication technologies (ICT) a term that stresses the role of unified communications and the integration of telecommunications (telephone lines and wireless signals), intelligent building management systems and audiovisual systems in modern information technology, including computer and network hardware and software. The term *ICT* is now also used to refer to the merging (convergence) of audiovisual and telephone networks with computer networks through a single cabling or link system.

Instant messaging (IM) a form of real-time direct text-based communication between two or more people using personal computers or other devices, along with shared clients. The user's text is conveyed over a network, such as the Internet. In many cases IM includes additional features. One broadly known type of IM is *Windows Live Messenger* (formerly named *MSN Messenger*), created by Microsoft, or *whatsapp*, an application created by Apple.

Internet relay chat (IRC) a form of real-time Internet text messaging (chat) or synchronous conferencing, mainly designed for group communication in discussion forums, but also allows one-to-one communication via private messages well as chat and data transfer.

Mobile telephone (mobile phone, cellular telephone, cell phone) electronic device used to make mobile telephone calls across a wide geographic area; it allows to make and receive telephone calls to and from the public telephone network which includes other mobiles and fixed-line phones across the world, by connecting to

a cellular network provided by a mobile network operator. In addition to telephony, modern mobile phones ("smartphones") also support a wide variety of other services such as text messaging, multimedia messaging, email, Internet access, short-range wireless communications (infrared, Bluetooth), business applications, gaming and photography.

Online social networks an online service, platform, or website that focuses on building and maintaining social networks or social relations among people. Social networking sites allow users to share ideas, activities, events, and interests within their individual networks. Currently, the most famous online social network is Facebook.

Short message service (SMS) text messaging, or texting, refers to the exchange of brief written text messages between fixed-line phone or mobile phone and fixed or portable devices over a network.

Voice over Internet protocol (VoIP) one of the family of Internet technologies, communication protocols, and transmission technologies for delivery of voice communications and multimedia sessions over Internet Protocol (IP) networks (Internet). Internet telephony refers to communications services – voice, fax, SMS, and/or voice-messaging applications – that are transported via the Internet, rather than the public switched telephone network.

Web 2.0 the term Web 2.0 is associated with web applications that facilitate participatory information sharing, interoperability, user-centered design and collaboration on the World Wide Web. A Web 2.0 site allows users to interact and collaborate with each other in a social media dialogue as creators of user-generated content, in contrast to websites where users are limited to the passive viewing of content that was created for them. Social networking sites or video sharing sites are typical examples of Web 2.0.

Further Reading

Alexa, R., Frank, M., Lester, D., 2010. An exploratory study of students' use of cell phones, texting, and social networking sites. Psychological Reports 107 (2), 402–404.

Bianchi, A., Phillips, J.G., 2005. Psychological predictors of problem mobile phone use. Cyberpsychology and Behavior 8 (1), 39–51.

Gardner, W., 2005. Just on click-sexual abuse of children and young people through the internet and mobile phone technology. Child Abuse Review 14, 448–449.

Huang, H., Leung, L., 2009. Instant messaging addiction among teenagers in China: shyness, alienation, and academic performance decrement. Cyberpsychology and Behavior 12 (6), 675–679.

Igarashi, T., Motoyoshi, T., Takai, J., Yoshida, T., 2008. No mobile, no life: self-perception and text-message dependency among Japanese high school students. Computers in Human Behavior 24 (5), 2311–2324.

Igarashi, T., Takai, J., Yoshida, T., 2005. Gender differences in social network development via mobile phone text messages: a longitudinal study. Journal of Social and Personal Relationships 22, 691–713.

Kamibeppu, K., Sugiura, H., 2005. Impact of the mobile phone on junior high-school students' friends hips in the Tokyo metropolitan area. Cyberpsychology and Behavior 8 (2), 121–130.

Madell, D., Muncer, S., 2004. Back from the beach but hanging on the telephone? English adolescents' attitudes and experiences of mobile phones and the internet. Cyberpsychology and Behavior 7 (3), 359–367.

Rakow, L.F., Navarro, V., 1993. Remote mothering and the parallel shift: women meet the cellular telephone. Critical Studies in Mass Communication 10, 144–157.

Reid, D.J., Reid, J.M., 2007. Text or talk? Social anxiety, loneliness, and divergent preferences for cell phone use. Cyberpsychology and Behavior 10 (3), 424–435.

Sánchez-Martínez, M., Otero, A., 2009. Factors associated with cell phone use in adolescents in the community of Madrid. Cyberpsychology and Behavior 12 (2), 131–137.

Srivastava, L., 2005. Mobile phones and the evolution of social behaviour. Behaviour and Information Technology 24, 111–129.

Takao, M., Takahashi, S., Kitamura, M., 2009. Addictive personality and problematic mobile phone use. Cyberpsychology and Behavior 12 (5), 501–507.

Walsh, S.P., White, K.M., Young, R.M., 2008. Over-connected? A qualitative exploration of the relationship between Australian youth and their mobile phone. Journal of Adolescence 31 (1), 77–92.

CHAPTER

92

Overuse of Social Networking

Enrique Echeburúa

University of the Basque Country UPV/EHU, Avda de Tolosa, 70 San Sebastián, Spain

OUTLINE

INTRODUCTION

A social networking site is an online place where a user can create a profile and build a personal network that connects him/her to other users. The social networking phenomenon has spread rapidly all over the world. Many Internet users have set up their own profile on social networking sites (e.g. Facebook, Twitter, or MySpace). Facebook, one of the main social networking sites, has over 500 million active users, with an additional 200 000 signing up each day. Facebook users visit the site on average two times a day and spend an average of 20 min per visit on the site. People online form relationships and social groups that provide emotional support and a sense of belonging. These groups can form intricate methods of communication, requirements for membership, and sets of standards and codes of conduct for their members.

One-third (35%) of American adult Internet users have a profile on an online social network site, 4 times more than 3 years ago, but still much lower than the 65% of online American teens who use social networks. Still, younger online adults are much more likely than their older counterparts to use social networks, with 75% of adults aged 18–24 using these networks, compared to just 7% of adults aged 65 and older. Overall, adults tend to use social networks for personal

http://dx.doi.org/10.1016/B978-0-12-398336-7.00092-9

reasons rather than professional. Despite comparatively lower levels of social network use, usage of social network sites by adults has increased markedly over the past 5 years. Demographics of social network users are shown in Table 92.1.

For girls, social networking sites are primarily places to reinforce pre-existing friendships; for boys, the networks also provide opportunities for flirting and making new friends. The question is if they are simply an easy way to be in close contact with acquaintances and friends or they are potentially addictive.

Despite the fact that the minimum age for most major social networking sites is usually 14 years, it is estimated that over a quarter of underage children have a profile on a social networking site because younger users often lie about their age when signing up for the website. However, content-generated risks from this new leisure activity have not been investigated in any detail.

Social networking websites allow teenagers to socialize and make friends with people they ordinarily would not approach and help shy people have an outlet for self-expression. They can write down their thoughts, add pictures of themselves, post messages for other people to read, and compile lists of their favorite interests. Users can design their personal homepages, adding different photographs, songs and even videos that can be played on the website alongside their personal information. Members can also link their profiles to those of other members and accumulate hundreds, even thousands, of online "friends," some of them from school or another social setting and some others whom they have never actually spoken to or met face-to-face.

All pleasurable behaviors can change our mood and consciousness and modify the brain chemistry. If they are used on a regular basis, they can negatively impact aspects of life functioning. The possibility of the user becoming addicted to the Internet increases tremendously due to the rewards given by accessing these networking sites. Socializing and dating on the net has become very popular recently.

That is, all addictive behaviors begin simply as pleasurable activities, which also serve to distract, connect to other friends and numb people from emotional discomfort. The problem is that what begins as a solution to social inhibition or to emotional trouble often produces secondary alterations in life functioning.

In sum, digital technologies are psychoactive and can alter mood states. Addictions of any type can be an exit or coping symptom of other ongoing wounds in the family or intrapersonal situation, so they may be amplifiers and sensitizers of previous problems.

TABLE 92.1 Demographics of Social Network Users

The percentage of online Americans in each demographic category who have a profile on a social network website:	
All adults	**35%**
Sex	
Men	35
Women	35
Age	
18–24	75*
25–34	57*
35–44	30*
45–54	19*
55–64	10
65+	7
Race	
White, non-Hispanic	31*
Black, non-Hispanic	43
Hispanic	48
Annual Household Income	
Less than $30 000	45*
$30 000–$49 999	38
$50 000–$74 999	30
$75 000+	31
Education	
Less than HS	43
HS grade	31
Some college	41*
College grade	33
Locale	
Urban	34*
Suburban	26
Rural	23

** These groups are significantly different from the other groups in the section.*
Source: Pew Internet & American Life Project Survey December 2008 Survey, n = 2253, with an n of 1650 Internet users. For Internet users the margin of error is ±3%.

OVERUSE OF SOCIAL NETWORKING SITES

Although Internet dependence is not equivalent to the serious problem of drug and alcohol addiction, there seem to be numerous commonalities in all addictions, specifically some alteration of mood and consciousness and some negative impact on a major life sphere such as health, family involvement, social relationships, or

work/school performance. All addictions produce initially pleasure, but they can also be an attempt to solve a problem (boredom/emptiness/anxiety). Specifically, the overuse of social networking serves to alter the mood and consciousness and therefore can generate a drug-like "high." In addition, there is preliminary evidence that Internet overuse can produce tolerance and withdrawal symptoms.

The major factors that seem to foster the overuse of the Internet and social networking are accessibility, availability, possibility of connecting to the world, accelerated intimacy, high stimulation, and perceived anonymity, associated to a sense of disinhibition. In some cases, teenagers and even elderly people feel self-confident, disclose their intimacy very easily and are careless about posting information that could become embarrassing in the future. In addition, many people experience an alteration of consciousness, lose track of time and space when they are online and are not critical about their own use. One thing the social networking sites provide to people is the ability to disconnect from the mundane of the daily grind and allow their mind to wander a little bit.

In the overuse of social networking, both the content and the Internet medium itself serve to accelerate their mutual impact and potency. Typically, the content consumed online is highly stimulating and reinforcing irrespective of the mode of access. However, the nature of the Internet medium itself often produces a more fertile context within which such stimulating content can be overused.

In some cases, parents can reinforce this abusive pattern in teenagers because as long as they have access to "their" technology they are silent and not disruptive to the family. However, overuse of computers becomes a reason for serious conflicts between parents and children.

In sum, in social networking sites people are looking for connection (pleasure) and distraction/numbing (avoidance of emotional discomfort). Probably the (over)use of social tools is more an issue of personal time-management that transcends the technology itself. However, in heavy social networking users, specifically in teenagers, the time invested in social media versus real-life interpersonal interaction can detract from real human contact and contribute to delayed and/or distorted social and emotional development.

EARLY WARNING SIGNS OF A POTENTIAL SOCIAL NETWORKING DEPENDENCE

There are not set rules for when a person can really become dependent on an online world. Anyway, there are some warning signs of overuse of social networking:

1. Increasing amount of time spent on social networking sites. More time is needed to achieve satisfaction and there are some difficulties tearing oneself away from the computer to deal with real life.
2. Using social networking sites while at work/school.
3. Preoccupation with social networking or withdrawal symptoms (irritability, lack of concentration, dysomnia) when not online.
4. Losing track of time and spending hours on social networking sites. Repeated unsuccessful efforts to limit or stop going online.
5. Feeling of annoyance if someone bothers him/her while he/she is online.
6. Using social networking to escape problems or gain relief from negative feelings.
7. Negative changes in school/work performance and psychosocial impairment. Important tasks and homework go undone.
8. Decreased social/familial involvement. Family and friends complain that he/she never seems to have time for them anymore. Previously enjoyed activities are abandoned.
9. Continued use despite adverse consequences. For example, he/she can outright lie when asked why he/she did not get a specific task completed or he/she can lose sleep because he/she stays online until late at night.
10. Changes in mood or in erratic or unpredictable behavior.
11. Increased secrecy around their digital behavior.

Sometimes overuse of social networking can mute the consciousness and people affected are not aware of the important tasks that are in to-be-done queue while consuming high-tech contents. Why? Because humans generally prefer the immediate pleasure over later; for example, studying to get grade A seems distant but one can enjoy spending one hour on Facebook or Twitter right now.

Internet heavy use can lead to more serious problems, such as mental illness, lying, lessened concentration, lower school grades, poor school attendance, dropping out of school, running away from home, and other family crises.

Internet dependence relates very closely with a gambling addiction and it may actually be considered as lack of impulse control. Several symptoms, such as hiding or lying about the amount of time spent online, anxiety while offline, and major impairment to offline functioning, bear close resemblance to problems faced by those addicted to gambling.

In sum, problems arise when users ignore family and work obligations because they find the social networking world a more enjoyable place to spend time than the real world.

IS OVERUSE OF SOCIAL NETWORKING A PRIMARY OR SECONDARY PROBLEM?

Hyper-networking is linked to a variety of concerns, including stress, depression, substance abuse, sleep troubles, and poor school/work performance. Actually teenagers who overuse social networking websites are far more likely to have engaged in risky behaviors such as smoking, drinking alcohol, or using drugs. Internet-enabled problems may cross many clinical areas. Mood disorders, impulsive disorders, and substance/alcohol abuse patterns should be assessed.

Overuse of social networking can be a manifestation of an underlying mental health disorder such as anxiety, depression or lack of impulse control. Users of social networking who feel lonely may look to the Internet as a way to communicate and express more easily their real selves without the fear of rejection and emotional strain. In these cases, Internet abuse is a way to cope with negative feelings and is secondary to other problems.

Narcissistic personalities having an overinflated sense of self can become heavy users of social networking sites. The need for these individuals to broadcast what they are doing, what they are thinking or what they have found or seen to an online audience that includes friends, colleagues, acquaintances, and even total strangers, is clearly indicative of narcissism or histrionism. They get in touch with hundreds of "friends" that are not real-life friends. In turn, histrionic people have a personality disorder marked by extreme emotional reactivity and an excessive need for attention and approval from others, so that they can overuse social networking as an outlet to their own problems. And the overuse of the Internet can reinforce these personality features.

However, when Internet heavy users are aware of their dependence and feel powerless to solve the problem by themselves, they can experience secondary problems, such as low self-esteem, depression, anxiety and social isolation, or aggravate previous disorders (emotional instability or social skills deficits). In this case, Internet abuse is a primary problem.

RISK FACTORS

Many people resort to social networking sites, but only a few become dependent on them. That means that some people are more vulnerable to Internet heavy use. Anyway, overuse of social networking sites, similar to other addictive behaviors, has multiple risk factors, and no single constellation of risk factors can alone predict if a particular problem will exist.

Adolescents are said to be most at risk for developing a dependence on the Internet. As they are familiar with the new technologies and they are in the process of maturing cognitively and developing their personality, they can become heavy Internet users more easily.

Personality Factors

Arousal and Sensation Seeking

Being permanently online acts as stimulant by providing an intensification of desirable emotions or sensations. Individuals who are underaroused, hypomanic, depressed, or get bored easily may find the excitement associated with social networking highly reinforcing. Connection to social working sites becomes a means of maintaining an optimal level of arousal. When such persons are underaroused or understimulated, the corrective effect of being online may be highly valued and sought. High-sensation seekers would then be more likely to be vulnerable to overuse social networking in order to maintain this optimal level of arousal and excitement.

Impulsivity

Impulsivity has consistently been associated with chemical and nonchemical addictions. Impulsivity means a tendency to get involved in rapid and unplanned reactions to stimuli before completing the processing of information, which means a decreased sensitivity to negative consequences of behavior. Impulsive people get involved in risky behaviors to ease tension or gain pleasure and do not pay attention to long-term consequences. An association between nonchemical addictions and attention deficit hyperactivity disorder (ADHD), a disorder related to poor psychosocial functioning in which impulsivity plays a major role, has often been reported.

Low Self-Esteem

Low self-esteem is a factor that the majority of social networking heavy users share in common. Those possessing low self-esteem are typically challenged in three key areas. They feel that they lack personal power and so their ability to influence others is compromised. Secondly, those with low self-esteem feel as if they were insignificant to others, lacking the affection and attention of others who hold them in low regard. Lastly, those possessing low self-esteem often hold themselves as incompetent in one or more relevant areas of life. They perceive

others as being more powerful and capable than they are. This negative feeling often results in ineffective communication and social conflicts, which further lead to diminished self-esteem. That is, users who lack social skills or suffer from low self-esteem are most likely to create a virtual person a online and use the Internet as a replacement for real-life socialization.

Those with low self-esteem often resort to social networking in an effort to show a different image of themselves, numb out the emotional discomfort and escape to a world that allows them a temporary release from their suffering and problems.

Social Skills Deficits

People with social skills deficits will turn to the Internet for social satisfaction since Internet communication is inherently less threatening due to increased anonymity and to non-face-to-face relationship. As a result, these users tend to be at higher risk for developing a heavy use of the Internet as it becomes the only medium in which they can fill their need for social interaction.

This preference for social networking is a direct result of the users' lack of confidence in their ability to communicate in face-to-face interactions. This low level of perceived self-worth in a social environment creates an inherent amount of stress, which leads people to seek out social support and manage their mood through methods available on the Internet.

Studies have struggled to ascertain whether addictive Internet use degrades a person's social skills to the point that Internet communication is the only viable source, or if the social skills were already lacking and therefore users turned to the net for an easier less threatening experience. Regardless, individuals who showed a preference for socialization on the Internet were significantly more likely to become addicted. Similarly, those users who were less satisfied with their real social life would be more likely to engage in problematic use of the Internet.

Mood States

Mood Regulation

Heavy social network use has been conceptualized as a negatively reinforced behavior through its ability to provide emotional relief from unpleasant stimuli. Dysphoric mood and thoughts prior to being online are an important trigger for abusing of social networking. However, it is not easy to know if these anxiety/depression symptoms precede or follow the onset of the overuse of social networking.

For some heavy users of the Internet, the urge to be online is associated with escaping from strong negative feelings such as sadness, loneliness, stress, or anxiety. Even though they know the overuse of social networking will create even stronger negative feelings in the long run, this knowledge is often put aside. The hope of a virtual relationship and the desire for a quick and easy escape are much stronger.

According to the studies in the field of alcohol/drugs abuse, women, when compared to men, are likely to use social networking sites for mood regulation and to help ease social relationships.

Dissociation

Certain activities performed on websites, among them gaming and social networking, have the potential to significantly alter normal functioning of the mind, which is especially dangerous for teenagers (and even young adults) whose minds are still growing.

Symptoms of dissociation involving altered experience of time and self have often been reported among people who abuse of social networking. Being online in these circumstances leads to subjective loss of the sense of time, depersonalization, escape, or dissociation (shift in personal identity and amnesiac episodes), with about 6% of individuals experiencing some significant impact on their lives. In these cases, there is a compulsion to dissociate from their real world and live in the social networking world instead.

Dissociation can be considered as an epiphenomenon of high states of arousal, which lead to narrowing of attention, escape, and a spectrum of increased cognitive distortions. Among heavy users of Internet, some may dissociate because of an imperative need to escape and may experience various levels of hypnotic states.

PROTECTIVE FACTORS

If only a few people become dependent on the social networking sites, that means there are lot of people using them in a positive way. And so, most recently, attention has begun to focus on the protective factors, which are thought to reduce the incidence of adolescent overuse of social networking. If many buffering factors are present, then behaviors like abuse of social networking are less likely under these conditions.

Looking at the individual, protective factors could include personal and social competence, optimism about the future, good problem-solving skills, and involvement in prosocial activities. Obviously this is not an exhaustive list. Another key element for an individual is resilience, which can be considered as the ability to cope with adversity in spite of a situation that one might not be able to change (e.g. living with an alcoholic parent). Some children are able to survive and thrive in difficult circumstances because their

individual strengths are dynamic, and so they become adapted and go on to develop in positive ways.

The foundations are laid in early childhood for later competence and resilience. Parents who are warm but structured with consistent rules and high expectations for behavior help children to develop prosocial behavior. The ability of children to control their emotions, attention, and behavior is a set of skills known as self-regulation. All of them help with the development of competence and resiliency which will buffer against risk factors.

Family cohesion, parenting attitude, and communication play a significant role as a protective factor. Children can grow up with a good self-esteem if parents praise student's achievements and accomplishments, acknowledge successes and abilities, and model a sense of optimism and a positive view of life. Such behaviors are supportive and tend to increase protective factors in children. In sum, families play an important role in preventing Internet abuse and must be considered when programs are developed to minimize excessive Internet usage by high-risk adolescents.

Risk and protective factors can affect youth at different stages of their lives. Early childhood risks, such as aggressive behavior, can be changed or prevented with family, school, and community interventions that focus on helping children develop appropriate behaviors. If not addressed, negative behaviors can lead to more risks, such as academic failure and social difficulties, which put children at further risk for later chemical or nonchemical addiction. Therefore, an important goal of prevention is to change the balance between risk and protective factors so that protective factors outweigh risk factors.

The key risk periods for all kinds of addictions are during major transitions in a person's life. The first big transition for children is when they leave the security of the family and enter school. Later, when they advance from elementary school, they often experience new academic and social situations, such as learning to get along with a wider group of peers. It is at this stage – early adolescence – that children are likely to encounter social networking sites for the first time.

The social environment of the school influences the healthy development of young people. Students who feel attached to their schools are less likely to engage in antisocial behavior or addictive practices. On the other hand, a feeling of alienation or not belonging can lead to behavior problems, substance use and antisocial activities.

The quality of the students' relationships with the teachers and their peers influences their sense of belonging. The protective effect of feelings of attachment is provided by warm relationships of mutual respect and by teachers who model positive interpersonal behavior, foster the normative value of helping and stimulate active student participation.

TREATMENT CONSIDERATIONS

Accurate Assessment

Many Internet heavy users refuse to admit they are addicted and do not seek treatment until major negative events take place. The majority of students will only seek help after failing out of school or having serious complications in their relationships. This lends support to the notion that early screening and detection is critical.

Before embarking on a treatment program, a sound assessment is required to inform the choice of treatment goal and content. Information should be attained on the evolution of heavy Internet use, family history, patterns of current use, degree of dependence, the extent of Internet dependence-related problems, reinforcement parameters maintaining the behavior, and the opportunities within the individual's environment for developing more adaptive responses.

The assessment process will assist the therapist in tailoring a treatment program to the needs of the individual, as well as elucidating on the most appropriate goals. According to the severity of dependence scale designed by Gossop and his colleagues to assess the degree of dependence experienced by users of different types of drugs, there are some relevant items in order to assess the severity of the Internet dependence in the last months:

1. Do you think your use of social networking sites is out of control?
2. Does the prospect of not being online make you anxious or worried?
3. Do you worry about your use of these sites?
4. Do you wish you could set a limit?
5. How difficult do you find it to set such a limit by yourself?

Motivational Enhancement Strategies for Change

The use of motivational enhancement strategies is a critical issue. Only when denial or defensive avoidance becomes untenable and a commitment to change is made, the patient is ready to treatment. Then, the individual begins to recognize the need for a change, gets ready for the change, and eventually works toward maintaining change.

During motivational interviewing people are encouraged to reach their own decision about change, while the role of the therapist is simply to facilitate this process

through clarification, advice when appropriate, accurate feedback, and empathy.

Unfortunately, there is no easy way to control the overuse of social networking. Being online has almost become an automatic behavior, like driving a car. That means that people will have to work hard at practising some new ways of thinking and behaving in order to stop the old habits from returning. Since hard work is not fun and learning new skills can be difficult, people need to increase their motivation for a new lifestyle. There are some motivational skills that may be useful for Internet heavy users to get involved in the process of change.

1. *People are able to control their behavior if they are willing to do it.* Some people manage not to be online for a few days and start feeling quite good, but when they cannot resist anymore the urge to be online, they feel miserable. They believe that they have no discipline, that they are impulsive, that they are out of control and, worst of all, that the goal of controlling the behavior is impossible for them. But this is a cognitive bias. It makes sense that people cannot stop when they do not have the skills that they need and should learn. It is not easy to give up an overlearned habit, but if people try again and again, it can be managed. Many people have given up in the past and more will give up in the future.
2. *People can calculate the amount of time they have spent online in the last weeks.* A good way to motivate people is to calculate the time they have wasted online in the last weeks, instead of being involved in interesting projects or in real social relationships. Anyway, even some of the heaviest users of Internet have been able to stop being online and rebuild their lives. There is always hope if they commit themselves to stopping.
3. *People should look back and look forward.* Another way to motivate people is to remember what their life was like before they started to have problems with being online. Heavy users of Internet are aware that their life was reasonably happy before being involved in an out-of-control use of social networking. Many heavy users of Internet often feel stressed, depressed, and anxious. The downward trend is very obvious. People can ask themselves whether the time has come to stop overusing social networking and arrest this trend. They can imagine the positive changes they can foresee for the next future.
4. *People should complete a decisional balance sheet.* The decisional balance sheet helps people to weigh up the pros and cons of overusing social networking. People must consider all of different areas of their life (school/work, family, social relationships, self-esteem, etc.) that have been affected by Internet dependence. As people can see, the number of disadvantages far outweighs the advantages. In addition, the advantages, such as excitement or dissociation, are often short-term, but disadvantages, such as poor school/work performance and loneliness, have a much longer-term impact.
5. *People should make a commitment from now on.* Stop overusing Internet should be the number one priority in life right now. That means that people are ready to attend the counseling sessions and practice the homework everyday. People need a positive self-talk about their ability to achieve the proposed goal.

At this point, people have admitted that they have a problem, they can not solve it by themselves, they have motivated themselves by weighing up the pros and the cons of their behavior and they have made a commitment to work on the problem.

Treatment Goals

The goal of treatment can focus on the attainment of controlled use of Internet and on the reduction of psychosocial problems either directly or indirectly related to the Internet dependence. Abstinence for ever may not be necessary. In the case of nonchemical addictions (except when referred to pathological gambling), it is possible for the individual to decide to cut back the overuse of social networking instead of deciding to quit for good.

However, the attainment of total abstinence from social networking sites by the individual during three or four weeks will be required before getting involved in a controlled use of Internet.

Treatments utilized depend upon several factors, including duration, intensity, and additional focus on stressors and comorbid problems. Thus, the clinician must determine whether hyper-networking is a primary or secondary problem. Counselors should evaluate psychological issues that might be impacting Internet-enabled behavior, as well as to examine major spheres of living (work/school, family, social relationships, etc.) that may be influencing quality of life and functioning. That is, the intervention offered should be tailored to the needs and circumstances of the individual.

In some cases, it is convenient to give support to the family system in order to reinforce the family bonds and to re-empower parents.

Cognitive-Behavioral Interventions

People who successfully overcome their overuse of social networking develop a variety of ways of coping.

Dealing with Urges

Overusers of social networking who decide to control the time spent online must learn to cope with temptations. The strongest urges usually take place during the first few weeks after quitting.

There are two types of effective coping. *Cognitive coping* uses self-talk to fight the urge to be online and consists of remembering past online problems and thinking about how well individuals are doing now to accomplish their goal. People naturally "talk to themselves" constantly. This self-talk can be positive and supportive or it can be negative and self-defeating. Positive self-talk can be the most immediate way to stop the urge and it helps make their behavior less automatic.

Behavioral coping means doing activities that will help people stay away from their abusive behavior. It is very important for individuals to make a change in their leisure and social activities. The key for breaking bad habits is to replace them by healthy ones. Having more leisure and social activities help individuals to distract until the urge passes, to cope with negative emotional states and to overcome loneliness. Thus, patients' self-confidence will improve and their life will become more balanced.

Limiting the Access to Computer

Individuals must be online in restricted periods of time and take them down in a notebook. It is convenient for them to set a limit, for example, 1 or 2 h a day, always at the same time. At the beginning of the treatment, after the period of total abstinence, people can resort to social networking in the living room, accompanied by other relatives (partner, parents, siblings, etc.). If the individual is getting on well with treatment, these requirements can be faded out later on.

Problem Solving and Emotion Management

Some people overuse social networking to escape from their problems, which can be many and various. Whatever the problem, the associated feelings are often so unpleasant that being online is seen as the best way to escape from everything for a while. Learning skills to deal effectively with these problems can help them to cut back being online.

Four-step problem solving (defining the problem, generating possible solutions, identifying the consequences of those solutions, choosing the best solution) is a very simple, yet highly effective, technique for dealing with problems. Sometimes people may not be able to solve some problems (a chronic illness, for example). If people are in this situation, they will need a way to cope with the distress caused by the problem.

Strong negative feelings, such as stress, anxiety, sadness, and loneliness, often trigger overuse of social networking episodes. For negative feelings to be dealt with adequately, they need to be acknowledged. Relaxation techniques can be an effective tool for controlling the urge to being online. Talking to a friend or family member can provide great relief. The other person can provide a fresh perspective on what is happening and can offer solutions that you may not have considered.

Telling Others of Your Plan

It is very helpful to tell those around individuals about their goal to cut back their overuse of social networking. Gaining support will help them share the burden.

In sum, the purpose of the whole programme is to rebuild patients' self-esteem, which is an important part of recovery.

Relapse Prevention

Relapse prevention programs aim to teach individuals who are trying to maintain changes in their behavior how to anticipate and cope with the risk situations.

A *lapse* is seen as being part of a transitional process, which may ultimately lead to favorable long-term outcome, while a *relapse* is a return to the original pattern of Internet overuse. Relapse is more likely in individuals who are not able to identify risk situations and have few coping resources.

As the duration of controlled use of social networking increases, the subject has the experience of coping effectively with one high-risk situation after another and the probability of relapse decreases accordingly. However, the failure to master a high-risk situation is likely to create a sense of powerlessness and hopelessness. If people slip, the severity of the overuse with each lapse is heavier and the consequences are more severe. The overuser may feel weak and out of control and lose the motivation to go on with the program.

Identifying High-Risk Situations

High-risk situations can be defined as those (places, feelings, thoughts, and events) where someone finds it very difficult to fight the urge and leads him/her to a non-moderate use of social networking sites. An abuse episode can be triggered after a period of moderate use when people unwittingly place themselves in high-risk situations, such as being online more than intended, to be locked in the room with the laptop, to be alone when online, or to be hooked to the social networking when they should have been sleeping.

High-risk situations can vary depending on the history of the individuals. Anyway, there are some common situations for everybody, such as social pressure, emotional trouble, or interpersonal conflicts. For

example, many overusers of social networking relapse after a period of moderate use because they are struggling with stress, sadness or loneliness. When they slip, their negative feelings grow stronger and it may be difficult to cope with them, resulting in further overuse.

People must learn from their slips by recognizing their triggers and risky situations. Each lapse can be a step closer to full recovery if slip is used as a learning experience to improve prevention strategies.

Coping Effectively with High-Risk Situations

Once triggers have been identified, people must find new ways of coping with them. The easiest coping mechanism for high-risk situations is to avoid them. This may include avoiding certain people who have a negative influence or avoiding locations (being locked in the bedroom with the laptop) where the problem is likely to occur. In some instances, avoidance is a good strategy. In other cases, avoidance is not appropriate because people need the laptop to work or get in touch with friends.

Rather, by being aware of this trigger, one can purposely engage in alternate activities during that time. Strategies for coping with unavoidable triggers may include discussion of feelings with a friend or relative; distraction, such as music, exercise, or engaging in a hobby; refocusing techniques, such as meditation, deep-breathing exercises or progressive muscle relaxation; and cognitive restructuring, such as positive affirmation statements (e.g. "I am worthwhile"), active problem solving, challenging the validity of negative thoughts, or guided imagery (imagining oneself in a different place or handling a situation appropriately).

Dealing with Other Life Problems

The onset of Internet abuse may be related to other life problems, such as emotional, family, social trouble, or substance abuse. These kinds of problems, whether related to non-moderate use of social networking, can make it difficult to overcome a serious dependence. People may realize in recovery that their Internet abuse was hiding or overshadowing other difficulties. Now that recovery is underway, people may be willing to take a closer look at these problems and work on them.

Developing a Balanced Lifestyle to Avoid Relapse

Besides being prepared for high-risk situations, relapse prevention programs also focus on general principles of mental health that, if followed, greatly reduce the likelihood of symptoms. These include factors such as balanced nutrition, regular exercise, sufficient sleep, health education, reciprocally caring relationships, productive and recreational interests, and spiritual development. Developing a social support system helps maintain a healthy lifestyle. This will make it much easier for vulnerable people to fight any temptation to lapsing.

CONCLUSION

Social networking has a multitude of uses, many of which are positive and beneficial. When overused, the Internet and its applications can lead to severely negative impacts for the user, both socially and personally. Although there are no statistics on "Internet addiction" – it is not an actual medical diagnosis – counselors say they are seeing more and more people who have crossed the line from social networking to social dysfunction.

People need to find a healthy balance between social networking and everyday life. The problem is when the virtual world is more important to people than their actual world. Young people can spend their after-school hours surfing MySpace and Facebook instead of participating in other activities. Social networking is an attempt to connect to others more efficiently, but perhaps with a cost of less depth and quality.

Overall, males undertake more risky behavior online than females and have more problems to set limits relating to the amount of time they spend online. Males also report more severe consequences in terms of effects on grades, jobs, and other performance related areas, as well as a higher likelihood of losing sleep, missing meals, and exercising less related directly to their amount of Internet usage. It is clear that unrestricted use of the social networking sites can be harmful, especially to males.

While alcoholics protect themselves from the addiction by not drinking anymore, it probably is not necessary to take potential social networking abuse to such extremes.

Computers and laptops located in private areas are likely to be a source of temptation when young people are alone. Teenagers who overuse social networking websites or texting are far more likely to have engaged in risky behaviors such as smoking, drinking alcohol, or using drugs. This overuse can have dangerous health effects on teenagers.

Future studies would enable better understanding of the detrimental effects that overuse of social networking may have on social satisfaction offline and on an overuser's life as a whole.

SEE ALSO

Internet: Immersive Virtual Worlds, The Cell Phone in the Twenty-first century: A Risk for Addiction or a Necessary Tool?

Glossary

Dissociation a psychological process involving alterations in identity or sense of self, which can include: a relatively mild and transient

sense that the world or the self are "unreal," more permanent states such as amnesia, and a disruption in the usually integrated functions of consciousness, memory, identity, and perception of the environment.

Impulsivity decreased sensitivity to the negative consequences of behavior; rapid, unplanned reactions to immediate stimuli before completing the information processing; and lack of regard for long-term consequences, often aimed at relieving tension or gaining pleasure.

Motivational enhancement therapy it focuses on repatterning the patient's behavior that is the result of ambiguous and undefined thoughts. This form of therapy is presented in a direct and patient-targeted manner that strives to transform undesired behaviors. Its goal is to help the patient clarify his or her own perceptions and beliefs in order to direct him or her in a more decisive way.

Protective factors factors in a person's life that promote mental health and well-being, such as a sense of attachment and belonging, supportive social networks (friends and family) or appropriate life skills (conflict resolution, anger management, and problem solving).

Relapse prevention cognitive-behavioral approach focused on the goal of identifying and coping effectively with high-risk situations that can trigger the relapse in the field of addictions.

Risk factors variable associated with an increased risk of having a problem. Risk factors are correlational and not necessarily causal, because correlation does not imply causation.

Sensation seeking strong need for new and varied experiences through uninhibited behavior and rejection of monotony, which may include engaging in risky activities.

Social networking site online place where a user can create a profile and build a personal network that connects him/her to other users.

Social networking-related harms negative consequences of becoming dependent on an online world (e.g. anxiety, depression, lying, social isolation, poor school grades, family crises).

Further Reading

Caplan, S.E., 2007. Relations among loneliness, social anxiety, and problematic internet use. Cyberpsychology and Behavior 10, 234–242.

Daley, D.C., Marlatt, G.A., 2006. Overcoming Your Alcohol or Drug Problem: Effective Recovery Strategies (therapist guide), second ed. Oxford University Press, Oxford.

Davidson, R.J., 1998. The treatment of substance abuse and dependence. In: Bellack, A.S., Hersen, M. (Eds.), Comprehensive Clinical Psychology, vol. 6. Elsevier, Amsterdam, pp. 567–585.

Eijnden, R., Spijkerman, R., Vermulst, A.A., Rooij, T.J., Engels, R., 2010. Compulsive internet use among adolescents: bidirectional parent–child relationships. Journal of Abnormal Child Psychology 38, 77–89.

Ellison, N.B., Steinfield, C., Lampe, C., 2007. The benefits of Facebook "friends": social capital and college students' use of online social network sites. Journal of Computer-Mediated Communication 12, 1143–1168.

Greenfield, D.N., 1999. Psychological characteristics of compulsive internet use: a preliminary analysis. CyberPsychology and Behavior 2, 403–412.

Griffiths, M., King, D., Delfabbro, P., 2009. Adolescent gambling-like experiences: are they a cause for concern? Education and Health 27, 27–30.

Hodgins, D.C., 2002. Becoming a Winner. Defeating Problem Gambling. Addictive Behaviors Laboratory, Calgary.

Milton, S., 2001. Stop Gambling. A Self-help Manual for Giving Up Gambling. Pan Macmillan, Sydney.

NIDA, 2003. Preventing Drug Abuse Among Children and Adolescents: A Research-based Guide for Parents, Educators, and Community Leaders, second ed. National Institute on Drug Abuse, Bethesda.

Prochaska, J.O., Norcross, J.C., Diclemente, C.O., 1994. Changing for Good: A Revolutionary Six-stage Program for Overcoming Bad Habits and Moving Your Life Positively Forward. Harper Collins, London.

Shaw, M., Black, D.W., 2008. Internet addiction. CNS Drugs 22, 353–365.

Valkenburg, P.M., Peter, J., 2007. Preadolescents' and adolescents' online communication and their closeness to friends. Developmental Psychology 65, 267–277.

Widyanto, L., Griffiths, M., 2006. "Internet addiction". A critical review. International Journal of Mental Health and Addiction 4, 31–51.

Young, K., Nabuco de Abreu, C., 2011. Internet Addiction: A Handbook and Guide to Evaluation and Treatment. John Wiley & Sons, Hoboken, NJ.

Index

Note: Page numbers with "*f*" denote figures; "*t*" tables

E

F

G

H

I

T